



PR 27-JUL-2000; 2000US-0221670P.  
 PR 07-AUG-2000; 2000US-0223170P.  
 PR 07-AUG-2000; 2000US-0223172P.  
 PR 07-AUG-2000; 2000US-0223460P.  
 PR 26-OCT-2000; 2000US-0244037P.  
 PR 26-OCT-2000; 2000US-0244111P.  
 PR 26-JUN-2001; 2001US-0301217P.  
 (DELT-) DELTAGEN INC.  
 PA Allen KD, Leviten MW;  
 XX MPI; 2002-154953/20.  
 DR P-PSDB; AAE23083.  
 PT Novel non-human transgenic animal, preferably transgenic mice comprising  
 PT disruption in target gene, e.g., trypsin gene, useful for identifying an  
 agent that modulates expression or function of target gene.  
 PS Example 2, Fig 2A, 74pp; English.  
 XX The present invention relates to non-human transgenic animals preferably  
 CC transgenic mice comprising disruption in target gene such as trypsin  
 CC gene. The invention also relates to compositions and methods relating to  
 CC the characterization of gene functions. The transgenic animals are useful  
 CC for identifying an agent that modulates the expression or function of a  
 CC target. They are useful for identifying an agent that modulates a  
 CC phenotype associated with a disruption in trypsin genes or trypsin  
 CC clotting factor protease-like genes by administering an agent to the  
 CC transgenic animal and determining whether the agent modulates the  
 CC phenotype where the agent has effect on decreased body weight, decreased  
 CC thymus weight, decreased thymus to body weight ratio, increased pre-pulse  
 CC inhibition, significant decrease in their response latency to the hot  
 CC plate test or a decreased response threshold to metrazol. Agents that  
 CC modulate the expression, function or activity of the target gene are  
 CC useful for treating a disorder associated with a mutation in trypsin  
 CC gene or in trypsin clotting factor protease-like gene. The transgenic  
 CC animals are useful for testing the efficacy of proposed genetic and  
 CC pharmacological therapies for human genetic diseases. They are useful as  
 CC models for diseases, disorders or conditions associated with phenotypes  
 CC relating to a disruption in a target and to identify drugs,  
 CC pharmaceuticals, therapies and interventions which may be effective in  
 CC treating a disease or other phenotypic characteristics of the animal. The  
 CC present sequence is epitithn gene. This sequence is used in the  
 CC exemplification of the invention  
 XX  
 SQ Sequence 3106 BP; 692 A; 862 C; 897 G; 655 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 3106; DB 6; Length 3106;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 3106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 301 ATTATCGAATGCGGGTTCAAAAAGTCTCAATGCGCATCGAATCAAAATGAGA 360  
 QY TCTTTTGTGATGCGTATGGAATCCACCTCCACAGATTATCAGCTGCGCAGCAG 420  
 Db 361 TCTTTTGTGATGCGTATGGAATCCACCTCCACAGATTATCAGCTGCGCAGCAG 420  
 QY TGAAGAGGCGCTGAAGCTGCTGTACATGAAAGTCTCTGCTGCTCTTACCAAGA 480  
 Db 421 TGAAGAGGCGCTGAAGCTGCTGTACATGAAAGTCTCTGCTGCTCTTACCAAGA 480  
 QY AGTGGCTGTAACTGCTTCACTGAGAGGAGTGTATCGGCTTCTACTGCTGAGTTCA 540  
 Db 481 AGTGGCTGTAACTGCTTCACTGAGAGGAGTGTATCGGCTTCTACTGCTGAGTTCA 540  
 QY AGTGGCTGTAACTGCTTCACTGAGAGGAGTGTATCGGCTTCTACTGCTGAGTTCA 540  
 Db 541 AGTGGCTGTAACTGCTTCACTGAGAGGAGTGTATCGGCTTCTACTGCTGAGTTCA 540  
 QY GCATCCCCCAGACCTGGCAGAGAGTTGATCGGCGCATAGGCTGTGAGACGATTGTA 600  
 Db 541 GCATCCCCCAGACCTGGCAGAGAGTTGATCGGCGCATAGGCTGTGAGACGATTGTA 600  
 QY CATTTGCCACCCCGAGAGACCGGCACTGAATCTTCTGCTTAACTGTGTGCTTCC 660  
 Db 601 CATTTGCCACCCCGAGAGACCGGCACTGAATCTTCTGCTTAACTGTGTGCTTCC 660  
 QY CCATTGACCCCAAGATGCTGCAAGAGACTGAGAGCAACAGTGAATTTGCTGATG 720  
 Db 661 CCATTGACCCCAAGATGCTGCAAGAGACTGAGAGCAACAGTGAATTTGCTGATG 720  
 QY CCATTGACCCCAAGATGCTGCAAGAGACTGAGAGCAACAGTGAATTTGCTGATG 720  
 Db 661 CCATTGACCCCAAGATGCTGCAAGAGACTGAGAGCAACAGTGAATTTGCTGATG 720  
 QY CCATTGACCCCAAGATGCTGCAAGAGACTGAGAGCAACAGTGAATTTGCTGATG 780  
 Db 721 CCATTGACCCCAAGATGCTGCAAGAGACTGAGAGCAACAGTGAATTTGCTGATG 780  
 QY CGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
 Db 781 CGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
 QY CGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
 Db 781 CGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
 QY TCCGAGCTTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 Db 841 TCCGAGCTTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 900  
 QY ATAGCTGAGCCCGCATGGAACCCACGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCT 960  
 Db 901 ATAGCTGAGCCCGCATGGAACCCACGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCT 960  
 QY CTTAACCTGACTTCT 1020  
 Db 961 CTTAACCTGACTTCT 1020  
 QY CTTAACCTGACTTCT 1020  
 Db 961 CTTAACCTGACTTCT 1020  
 QY GTGAGCGGCTTTTGAAGTGAACCAAGAGGACATTTAGAGGCTTATTCAGGCTCACT 1140  
 Db 1081 GTGAGCGGCTTTTGAAGTGAACCAAGAGGACATTTAGAGGCTTATTCAGGCTCACT 1140  
 QY GTGAGCGGCTTTTGAAGTGAACCAAGAGGACATTTAGAGGCTTATTCAGGCTCACT 1140  
 Db 1081 GTGAGCGGCTTTTGAAGTGAACCAAGAGGACATTTAGAGGCTTATTCAGGCTCACT 1140  
 QY ACCCGCCCAATCAATCTGCAATGGAATATCAAGTGTCCCAACCAACCGGAACGTAAGG 1200  
 Db 1141 ACCCGCCCAATCAATCTGCAATGGAATATCAAGTGTCCCAACCAACCGGAACGTAAGG 1200  
 QY ACCCGCCCAATCAATCTGCAATGGAATATCAAGTGTCCCAACCAACCGGAACGTAAGG 1200  
 Db 1141 ACCCGCCCAATCAATCTGCAATGGAATATCAAGTGTCCCAACCAACCGGAACGTAAGG 1200  
 QY TGGGCTTGAATCTTCTTATCTGCTGAGGACCCCAAGTACAGTGGGCTCTGCAACAAAG 1260  
 Db 1201 TGGGCTTGAATCTTCTTATCTGCTGAGGACCCCAAGTACAGTGGGCTCTGCAACAAAG 1260  
 QY TGGGCTTGAATCTTCTTATCTGCTGAGGACCCCAAGTACAGTGGGCTCTGCAACAAAG 1260  
 Db 1201 TGGGCTTGAATCTTCTTATCTGCTGAGGACCCCAAGTACAGTGGGCTCTGCAACAAAG 1260  
 QY ACTATGTGAAGATCAACCGGGAAGAGTACTGCGGTGAGAGGCTCCAGTTTGTGTAGAGA 1320  
 Db 1261 ACTATGTGAAGATCAACCGGGAAGAGTACTGCGGTGAGAGGCTCCAGTTTGTGTAGAGA 1320  
 QY GCAACAGCAGCAAGATTAAGTCACTTCTGATCTGATCTGATCAACAGCAACCGGCT 1380  
 Db 1321 GCAACAGCAGCAAGATTAAGTCACTTCTGATCTGATCTGATCAACAGCAACCGGCT 1380  
 QY TCTTATGTGAAGATCAACCGGGAAGAGTACTGCGGTGAGAGGCTCCAGTTTGTGTAGAGA 1440  
 Db 1381 TCTTATGTGAAGATCAACCGGGAAGAGTACTGCGGTGAGAGGCTCCAGTTTGTGTAGAGA 1440

Db 1381 TCCTAGCTGAGTACCTCTCCAGACTCCACAGACCCGTCGCCAGGATGTTCAATGCA 1440  
 QY 1441 AGACTGAGACGGTGATCCGAAAGAACTGCGCTGCGACGCGCTGGGAGACTGCCGAGT 1500  
 Db 1441 AGACTGAGACGGTGATCCGAAAGAACTGCGCTGCGACGCGCTGGGAGACTGCCGAGT 1500  
 QY 1501 ATAGTATGAGCGCTTACTGCGAGATGCAATGCCACACCACTGTTACCTGCAAAAACCA 1560  
 Db 1501 ATAGTATGAGCGCTTACTGCGAGATGCAATGCCACACCACTGTTACCTGCAAAAACCA 1560  
 QY 1561 TCTGCAAGCCCTCTTCTGCGGTCTGTGACAGTGTCAACAGACTGTGGGAGCGAAGTAGC 1620  
 Db 1561 TCTGCAAGCCCTCTTCTGCGGTCTGTGACAGTGTCAACAGACTGTGGGAGCGAAGTAGC 1620  
 QY 1621 AGAGGCGCTGACAGTGTCTGCTGGAGATTCAAGTGTTCCAAATGGGAAGTGTCTCCCTC 1680  
 Db 1621 AGAGGCGCTGACAGTGTCTGCTGGAGATTCAAGTGTTCCAAATGGGAAGTGTCTCCCTC 1680  
 QY 1681 AGAGCCAGAGTGTATGGAAGAGCAACTGTGAGATGGGTCTGACAGGCTTCAATGTG 1740  
 Db 1681 AGAGCCAGAGTGTATGGAAGAGCAACTGTGAGATGGGTCTGACAGGCTTCAATGTG 1740  
 QY 1741 ACAGCGTGAATGTCTCTCTTGACCAAAATATACCTACCGCTGCCAAAATGCGCTCTGTC 1800  
 Db 1741 ACAGCGTGAATGTCTCTCTTGACCAAAATATACCTACCGCTGCCAAAATGCGCTCTGTC 1800  
 QY 1801 TGAGCAAGGCAACCTCTGAGTGTGATGGAGAGCGAGCTGTAGCGATGGCTCCGATGAG 1860  
 Db 1801 TGAGCAAGGCAACCTCTGAGTGTGATGGAGAGCGAGCTGTAGCGATGGCTCCGATGAG 1860  
 QY 1861 AAAACGTGACTGTGGGCTGGCATCTCTTAAACAAAGAGCTGGCGTGGTGGTGACAGA 1920  
 Db 1861 AAAACGTGACTGTGGGCTGGCATCTCTTAAACAAAGAGCTGGCGTGGTGGTGACAGA 1920  
 QY 1921 ATGCGAGCAGAGGCGAGTGGCCCTGCGAGAGTGAAGCTCCACGCGCTGGGCAAGGCACT 1980  
 Db 1921 ATGCGAGCAGAGGCGAGTGGCCCTGCGAGAGTGAAGCTCCACGCGCTGGGCAAGGCACT 1980  
 QY 1981 TGTGTGGGCGCTGTGCTCATCTCTCTGACTGTGTGTCTGTGAGCTCATTTGCTTTCAG 2040  
 Db 1981 TGTGTGGGCGCTGTGCTCATCTCTCTGACTGTGTGTCTGTGAGCTCATTTGCTTTCAG 2040  
 QY 2041 ATGACAAAATTTCAAGTACTGAGTACTAGACTAGAGTGTGAGCGGCTTCTGCGGTCTGCTG 2100  
 Db 2041 ATGACAAAATTTCAAGTACTGAGTACTAGACTAGAGTGTGAGCGGCTTCTGCGGTCTGCTG 2100  
 QY 2101 ACCAGAGCAGAGCGAGTGTCTGCGGAGTCAAGAGCTGAAGCTCAAAAGTATCATCACCC 2160  
 Db 2101 ACCAGAGCAGAGCGAGTGTCTGCGGAGTCAAGAGCTGAAGCTCAAAAGTATCATCACCC 2160  
 QY 2161 ACCCTTCTCTTCAATGATTTCACTTGTGACTATGACATGCGCTTGTGAGCTGAGAAAT 2220  
 Db 2161 ACCCTTCTCTTCAATGATTTCACTTGTGACTATGACATGCGCTTGTGAGCTGAGAAAT 2220  
 QY 2221 CGGTGAGTACAGCAACCGTGTGGCGCCCATGTGCTGCTGCTGATGCTAACCTATGTTCC 2280  
 Db 2221 CGGTGAGTACAGCAACCGTGTGGCGCCCATGTGCTGCTGCTGATGCTAACCTATGTTCC 2280  
 QY 2281 CTGCTGAGAGCGCATCTGAGTCAAGAGCTGCGGAGCACAAAAGAGGAGTACCGAG 2340  
 Db 2281 CTGCTGAGAGCGCATCTGAGTCAAGAGCTGCGGAGCACAAAAGAGGAGTACCGAG 2340  
 QY 2341 CGCTATCTCTGAGAAAGGCTGATCCGTGTATCAACCAAGCACTGTGAGAGCTCA 2400  
 Db 2341 CGCTATCTCTGAGAAAGGCTGATCCGTGTATCAACCAAGCACTGTGAGAGCTCA 2400  
 QY 2401 TGCCGAGAGATCAACCCACGATGATGTGTGTGTTTCTCAGTGGGGTGTGAGT 2460  
 Db 2401 TGCCGAGAGATCAACCCACGATGATGTGTGTGTTTCTCAGTGGGGTGTGAGT 2460  
 QY 2461 CTTGCAAGGCTGACTCTGTGTGCGCTTGTCAAGGCGGAGAAAGATGGGCAATGTTCC 2520  
 Db 2461 CTTGCAAGGCTGACTCTGTGTGCGCTTGTCAAGGCGGAGAAAGATGGGCAATGTTCC 2520

QY 2521 AGGCTGTGTGTGAGCTGGGGTGAAGGCTGCGCTGAGAGAACAGGCGAGGCTGTACA 2580  
 Db 2521 AGGCTGTGTGTGAGCTGGGGTGAAGGCTGCGCTGAGAGAACAGGCGAGGCTGTACA 2580  
 QY 2581 CAAGGCTCCCTGTATGTTTGGGACTGATCAAGAGCACTGGGGTATATACAGATGAC 2640  
 Db 2581 CAAGGCTCCCTGTATGTTTGGGACTGATCAAGAGCACTGGGGTATATACAGATGAC 2640  
 QY 2641 AGACAGCGGACCAAAACACCAAGGATGCGCCGACATGACACCTGATACAGAGAG 2700  
 Db 2641 AGACAGCGGACCAAAACACCAAGGATGCGCCGACATGACACCTGATACAGAGAG 2700  
 QY 2701 GAACACTGACGACATTTATGCTGTGCGTCCCGCCCAACACACCACTGATGACT 2760  
 Db 2701 GAACACTGACGACATTTATGCTGTGCGTCCCGCCCAACACACCACTGATGACT 2760  
 QY 2761 GCATCTTATGAGACTCAGAGTTCTTCCAAAGTGGGACCCCTCAAGAGTTGAGAGAGAC 2820  
 Db 2761 GCATCTTATGAGACTCAGAGTTCTTCCAAAGTGGGACCCCTCAAGAGTTGAGAGAGAC 2820  
 QY 2821 TTGCGTGTAGCGGCGCCAGCTTGGGGCAAGGCTTGTATGACAGCTTCCCTTAAGC 2880  
 Db 2821 TTGCGTGTAGCGGCGCCAGCTTGGGGCAAGGCTTGTATGACAGCTTCCCTTAAGC 2880  
 QY 2881 CTGAGCTGGGTGAAGATGATGTCTGCTCCGAGAGCTGTCCAACTGTCAATGAGCTCCC 2940  
 Db 2881 CTGAGCTGGGTGAAGATGATGTCTGCTCCGAGAGCTGTCCAACTGTCAATGAGCTCCC 2940  
 QY 2941 GGGAGCCTATGAGAGAGGAGGCTCAGAGGTCACTCTTTCAAGAAAGCGGCACTTACGA 3000  
 Db 2941 GGGAGCCTATGAGAGAGGAGGCTCAGAGGTCACTCTTTCAAGAAAGCGGCACTTACGA 3000  
 QY 3001 ACCCCAGAAAGAGTGTACTTAAGCTGAAATTTGTTTGTCTGCTGCTGAGGCTGTAT 3060  
 Db 3001 ACCCCAGAAAGAGTGTACTTAAGCTGAAATTTGTTTGTCTGCTGCTGAGGCTGTAT 3060  
 QY 3061 TTGAGAGTAAACATTTTATTTCTTTTAAAAA 3106  
 Db 3061 TTGAGAGTAAACATTTTATTTCTTTTAAAAA 3106  
 RESULT 2  
 AAA88493  
 ID AAA88493 standard; cDNA; 3149 BP.  
 XX  
 AC AAA88493;  
 XX  
 DT 22-JAN-2001 (first entry)  
 XX  
 DE Human matrixase cDNA.  
 XX  
 KW Matrixase; serine protease; human; breast cancer; pre-malignancy;  
 KW actinic keratosis; leukoplakia; Barrett's epithelium;  
 KW colunar metaplasia; ulcerative colitis; bowenoid papulosis;  
 KW adenomatous colorectal polyp; Oyerat erythroplasia;  
 KW vulvar intraepithelial neoplasia; tumour; metastasis; therapy; ss.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 36..2603  
 FT FT /\*tag= c  
 FT PN  
 XX  
 PD 14-SEP-2000.  
 XX  
 PF 10-MAR-2000; 2000WO-US006111.  
 XX  
 PR 12-MAR-1999; 99US-012406P.  
 XX  
 PA (GBOU ) UNIV GEORGETOWN.

XX Dickson RB, Lin C, Johnson M, Wang S, Emyedy I;  
XX WPI; 2000-594268/56.  
XX P-PSDB; AAB19552.

XX Treating malignancies, premalignant and pathologic conditions in a  
XX subject, comprises administering matrixinase modulating agent.

XX Claim 9; Fig 15; 116pp; English.

XX The present sequence is that of cDNA coding for the full-length form (see  
XX AAB19521) of human matrixinase, a trypsin-like protease. cDNA encoding a  
XX truncated form of matrixinase is given in AAB8492. Either form can be  
XX incorporated into a vector and used in a claimed method for making  
XX recombinant matrixinase. The zymogen (inactive) form of matrixinase is a  
XX single-chain protein. The active 2-chain form strongly interacts with  
XX fragments of a Kunitz-type serine protease inhibitor (hepatocyte growth  
XX factor activator inhibitor, HAI-1) to form SDS-stable complexes. In  
XX breast cancer cells, matrixinase is present mainly as the uncomplexed  
XX form. Only the complexed matrixinase is detected in human milk. The  
XX invention is directed to a method of detecting a malignancy or a pre-  
XX malignant lesion in breast or other tissue by detecting the presence of  
XX single- or 2-chain forms of matrixinase in the tissue. The object is to  
XX inhibit tumour onset, tumour growth and metastasis. Malignancies and pre-  
XX malignant conditions characterised by expression of the zymogen or  
XX activated form of matrixinase are treated by administering an inhibitor of  
XX matrixinase, especially a Bowman-Birk inhibitor. The pre-malignant  
XX condition is atypical ductal hyperplasia of the breast, atypical  
XX keratosis, leukoplakia, Barrett's epithelium of the esophagus, ulcerative  
XX colitis, adenomatous colorectal polyps, erythroplasia of the Queyrat,  
XX Bowen's disease, Bowenoid papulosis, vulvar intraepithelial neoplasia or  
XX dysplastic changes to the cervix. The invention also provides methods for  
XX in vivo or in vitro diagnosis of malignancy or pre-malignant lesion, and  
XX methods of identifying matrixinase modulators, including activators and  
XX inhibitors

XX Sequence 3149 BP; 647 A; 959 C; 962 G; 581 T; 0 U; 0 Other;

XX Query Match 60.84; Score 1889.8; DB 3; Length 3149;

XX Best Local Similarity 81.48; Pred. No. 0;  
XX Matches 2222; Conservative 0; Mismatches 504; Indels 5; Gaps 3;

XX 50 GACCGCAAAACCATGGGTAGCATCGGGCGCGCAAGGCGGAGGGCTTCAGAGACTT 109  
XX 23 GGCTCGGGGACCATGGGAGCGATCGGCGCGCAAGGCGGAGGGCGGAGAGACTT 82  
XX 110 CGGGCGGGGACTCAAGTCAACATCCCGCTAGAGAAATGAATGGCTTTGAGAGGGTGT 169  
XX 83 CGGGCGGGGACTCAAGTCAACATCCCGCGCAAGAAATGAATGGCTTTGAGAGAGGGT 142  
XX 170 GGAAGTCTCTGCTCGGAGCAATGCCAAGAAAGTGAAGAGGAGGCGCGAGGGT 229  
XX 143 GGAAGTCTCTGCTCGGAGCAACATCGTCAAGAAAGTGAAGAGGCGCGAGGGT 202  
XX 230 GGTGCTGGTGAAGAGTGTTCAGCTTCTTGTCTCTCCCTCAATGGCTTGTGCTGT 289  
XX 203 GGTGCTGGTGAAGAGTGTTCAGCTTCTTGTCTCTTGTCTTGTCTGGGAGTCCGCTTCT 262  
XX 290 GTGGCACTTCATTAATCGAAATGTGCGGGTTCAAAAGTCTTCAATGGCCATTTGAGAT 349  
XX 263 GTGGCACTTCATTAATCGAAATGTGCGGGTTCAAAAGTCTTCAATGGCCATTTGAGAT 322  
XX 350 CACAAATGAGATCTTTCTGATGCTATGAGAACTCCCACTCCAGAGTTTATGAGCT 409  
XX 323 CACAAATGAGATCTTTCTGATGCTATGAGAACTCCCACTCCAGAGTTTATGAGCT 382  
XX 410 GGGCAGCAGAGTGAAGAGAGCGCTGAGAGTGTATGATGAATGAAGTCCCTGTCTGCTGCT 469  
XX 383 GGGCAGCAGAGTGAAGAGAGCGCTGAGAGTGTATGATGAATGAAGTCCCTGTCTGCTGCT 442  
XX 470 CTACCAAGAAGTGGCTGTAACTGCTTCAATGAGGAGCAATGCTATGCTTACTGCT 529

DB 443 CTACCAAGAAGTGGCTGTAACTGCTTCAATGAGGAGCAATGCTATGCTTACTGCT 502  
QY 530 GTGAGAGTTGAGATCTCCCACTCCCACTGGGAGAGAGTGTATGCGCCATGCTGTGGA 589  
DB 503 GTTGAAGTTGAGATCTCCCACTCCCACTGGGAGAGAGTGTATGCGCCATGCTGTGGA 562  
QY 590 GCGAGTTGATCAATTCGCAATCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 649  
DB 563 GCGAGTATGATCTGCTCCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 622  
QY 650 GGTGGCTTCCCACTGAG 709  
DB 623 GGTGGCTTCCCACTGAG 682  
QY 710 TGGCTGAG 769  
DB 683 TGGCTGAG 742  
QY 770 TGGCTGAG 829  
DB 743 CCGCTACCGGCTGATGCGCTGCGAGTGGGCGCTGCGGAGAGAGAGAGAGAGAGAG 802  
QY 830 GAGCTTCACTTCGAG 889  
DB 803 GAGCTTCACTTCGAG 862  
QY 890 CACGCTGATGATAGCTGAG 949  
DB 863 GAGGCTTCACTTCGAG 922  
QY 950 CTGCTCACTTCGAG 1009  
DB 923 CTGCTCACTTCGAG 982  
QY 1010 GATTAACCAATGAG 1069  
DB 983 GATTAACCAATGAG 1042  
QY 1070 GATGAG 1129  
DB 1043 GATGAG 1102  
QY 1130 TCCAG 1189  
DB 1103 CCGAG 1162  
QY 1190 GAACTGAG 1249  
DB 1163 GAACTGAG 1222  
QY 1250 CTGAG 1309  
DB 1223 CTGAG 1282  
QY 1310 TGTGAG 1369  
DB 1283 GGTGAG 1342  
QY 1370 GGAACAG 1429  
DB 1343 GGAACAG 1402  
QY 1430 GTTCAATGAG 1489  
DB 1403 GTTCAATGAG 1462  
QY 1490 CTGAG 1549  
DB 1463 CTGAG 1522  
QY 1550 CAAAAG 1609  
DB 1523 CAAAAG 1582

QY 1610 CGGAAGTACAGAGAGGCTGACAGCTGTCTGCTGAGAGTTTCAAGTGTTCATATGAGAA 1669  
 Db 1583 CAACAGGACGAGAGGAGGCTGACAGTGTCTGCGCCCAAGACCTTGAGTGTTCATATGAGAA 1642  
 QY 1670 GTGTCTCCCTCAGAGAGGAGTAAATGGAAGAGACAATGAGAGATGGGTGTGCGA 1729  
 Db 1643 GTGTCTCCCAAAAAGCCAGAGTGTCAATGGAAGAGAGAGCTGTGGGAGCGGTCCGCGA 1702  
 QY 1730 GGCCTTCAATGAGACAGGCTGAATGTCTGTGACCAAAATATACCTACCGCTGCCAAA 1789  
 Db 1703 GGCCTCTGTGCCCCAAGGTGAACGTCTGTCTGTACCAAAACAACCTACCGCTGTCTCAA 1762  
 QY 1790 TGGCTCTGTCTGAGCAAGGCGCAACCTGATGTGATGAGAGACGAGCTGTAGCCATG 1849  
 Db 1763 TGGCTCTGTCTGAGCAAGGCGCAACCTGATGTGATGAGAGAGAGAGCTGTAGCCAG 1822  
 QY 1850 CTCGATGAGAAAACCTGTGACCTGTGGGCTGCGATCTTTTACCAACAGGCTGCGTGT 1909  
 Db 1823 CTCGATGAGAAAACCTGTGACCTGTGGGCTGCGATCTTTTACCAACAGGCTGCGTGT 1882  
 QY 1910 TGGTGGACAGAAATGCGAGCGAGCGAGTGGCCCTGAGAGTGGCTCCACGCTGGG 1969  
 Db 1883 TGGGAGGACAGAGATGCGAGATGAGGCGAGTGGCCCTGAGAGTGGCTCCACGCTGGG 1942  
 QY 1970 CCAAGGCGCACTTGTGTGGGCGCTGCGTCACTCTCTGACTGGTGGTCTGTGCACTGA 2029  
 Db 1943 CCAAGGCGCACTTGTGTGGGCGCTGCGTCACTCTCTGACTGGTGGTCTGTGCACTGA 2002  
 QY 2030 TTGCTTTGAGATGACAAAATTTCAAGTACTAGACTACAGACTAGTGAACGCGCTTCT 2089  
 Db 2003 CTGCTACATGATGATGACAGAGATTTCAAGTACTAGACTACAGACTAGTGAACGCGCTTCT 2062  
 QY 2090 GGGTCTGTGAGCAAGAGCAAGCGAGTGTCTGTGGGCTGAGAGAGCTGAAGCTCAAG 2149  
 Db 2063 GGGCTTGCACGACAGAGCAAGCGAGTGTGTGGGCTGAGAGAGAGCTGAAGCTCAAG 2122  
 QY 2150 TATCATCACCAACCTTCTCTTCAATGATTTCACTTGTGACTATGACATGCGCTTGTG 2209  
 Db 2123 CATCATCTCCCAACCTTCTCTTCAATGATTTCACTTGTGACTATGACATGCGCTTGTG 2182  
 QY 2210 GCTGAGAAAGTGGAGTGAAGACGAGCGTGGTGGCGCCCACTGTGCTGCTGATGCTAC 2269  
 Db 2183 GCTGAGAAAGTGGAGTGAAGACGAGCGTGGTGGCGCCCACTGTGCTGCTGATGCTAC 2242  
 QY 2270 CCATGTCTTCTGCTGTGAGAGCGCATCTGGGTCAAGAGCTGGGCGACACAAAAGAGG 2329  
 Db 2243 CCATGTCTTCTGCTGTGAGAGCGCATCTGGGTCAAGAGCTGGGCGACACAAAAGAGG 2302  
 QY 2330 AGGTACCGGAGCGCTGATCTTGCAGAGGAGTGAATCGGTGATCAACAGACCACTG 2389  
 Db 2303 AGGTACCGGAGCGCTGATCTTGCAGAGGAGTGAATCGGTGATCAACAGACCACTG 2362  
 QY 2390 TGAGAGACCTCATGCGCGAGCAGATCAACCCAGAAATGATGTGGGTCTTCTCAATG 2449  
 Db 2363 TGAGAGACCTCATGCGCGAGCAGATCAACCCAGAAATGATGTGGGTCTTCTCAATG 2422  
 QY 2450 GGGTGTGAGCTCTTCCAGAGGAGTACTGTGGTGGCCCTTGTCAACCGCGAGAAAGATG 2509  
 Db 2423 GGGTGTGAGCTCTTCCAGAGGAGTACTGTGGTGGCCCTTGTCAACCGCGAGAAAGATG 2482  
 QY 2510 GCGAATGTTTCCAGAGTGTGTGTGATGCTGGGAGTGAAGGCTGTGAGAGAACAGCC 2569  
 Db 2483 GCGAATGTTTCCAGAGTGTGTGTGATGCTGGGAGTGAAGGCTGTGAGAGAACAGCC 2542  
 QY 2570 AGGCGTGTACACAGAGGCTCCCTGTAGTTCGAGAGCTGTGATCAAAAGAGACACTGGGAT 2629  
 Db 2543 AGGCGTGTACACAGAGGCTCCCTGTAGTTCGAGAGCTGTGATCAAAAGAGACACTGGGAT 2602  
 QY 2630 GCAAGTGTACACAGAGGCTCCCTGTAGTTCGAGAGCTGTGATCAAAAGAGACACTGGGAT 2689  
 Db 2603 GCAAGTGTACACAGAGGCTCCCTGTAGTTCGAGAGCTGTGATCAAAAGAGACACTGGGAT 2659

QY 2690 ATACAGAGAGAGAACATGACAGCAATTTATGCTGTGGCTCCCGCCCAACACACCA 2749  
 Db 2660 GTGCAGG-CCTGAGAGGTGAGAGACTGAGCCGCTGACTGACACAGCGCC-CCAGAACATA 2717  
 QY 2750 GACTGTGAACTGCATCTTGAAGACTCAGAGT 2780  
 Db 2718 CACTGTGAAGTCAATCTCCAGGAGCTCCAAAT 2748  
 RESULT 3  
 AAF28099  
 ID AAF28099 standard; cDNA; 3142 BP.  
 XX  
 AC AAF28099;  
 XX  
 DT 06-JUN-2001 (first entry)  
 XX  
 DE Human membrane-type serine protease MT-SPI coding sequence.  
 XX  
 KW Human; membrane-type serine protease; MT-SPI; cancer; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT CDS Location/Qualifiers  
 FT 37..2602  
 FT /\*tag= a  
 FT /product= "MT-SPI"  
 XX  
 PN MO200123524-A2.  
 XX  
 PD 05-APR-2001.  
 XX  
 PF 02-OCT-2000; 2000MO-US027250.  
 XX  
 PR 30-SEP-1999; 99US-00410362.  
 XX  
 PA (REGC ) UNIV CALIFORNIA.  
 XX  
 PI Craik GS, Takeuchi T, Shuman M;  
 XX  
 DR WPI, 2001-245002/25.  
 DR P-PSDB; AAB35465.  
 XX  
 PT New nucleic acid encoding a membrane type serine protease, useful for the  
 PT diagnosis, prognosis and treatment of cancer, particularly metastatic  
 PT cancers.  
 XX  
 PS Claim 4; Fig 1; 102pp; English.  
 XX  
 CC The present invention provides the protein and coding sequences for the  
 CC novel human membrane-type serine protease MT-SPI. Increased expression of  
 CC this protein is associated with cancer, and so the sequences can be used  
 CC in cancer diagnosis and the identification of treatments. The present  
 CC sequence is the MT-SPI coding sequence  
 XX  
 SQ Sequence 3142 BP; 635 A; 958 C; 965 G; 584 T; 0 U; 0 Other;  
 Query Match 60.7%; Score 1884.8; DB 4; Length 3142;  
 Best Local Similarity 82.8%; Pred. No. 0;  
 Matches 2164; Conservative 0; Mismatches 447; Indels 1; Gaps 1;  
 QY 34 CTGAGACCGGAGTGCAGCCGCAAAACCATGGGTGAGCAATCGGGGCGCAAGCGCGAG 93  
 Db 9 CTGAGACCGGAGGAGCGG-CCTGAGGAGCAATGGGAGCGATCGGCCCGCAAGGCGGAG 67  
 QY 94 GGGGCTTCAAGACTTGTGGCGCGAGACTCAAGTCAACTCCCGGCTAGAGAACATGATG 153  
 Db 68 GGGGCGCGAAGACTTGTGGCGCGAGACTCAAGTCAACTCCCGGCAAGAAAGTGAATG 127  
 QY 154 GCTTGGAGAGAGGTGTGAGTCTCTGCTGTGAGAACATGCGCAAGAAATGTGAGAGCAG 213  
 Db 128 GCTTGGAGAGAGGTGTGAGTCTCTGCTGTGAGAACATGCGCAAGAAAGTGTGAGAGCAG 187

[illegible]

D	b	1268	GAGAGAGTCCCAAGTTCGTGTCACACAGAAACAGAAACAAGATCAAGTTCCGCTTCCACT	1327
Q	y	1254	CTGATCACTCGTACACGAGACACCGGGTCTCTAGCTGAGTAACTCTCTCCAGATCCAAAG	1413
D	b	1328	CAGATCACTCTTACACCGACACCGGCTCTTAGCTGAATAACTCTCTACGATCCAGTG	1387
Q	y	1414	ACCGTGGCCAGGGATGTTTCATGTCACAAAGCTGACGTGATCCGAAAGAACTCGCT	1473
D	b	1388	ACCATGCCCCGGGGAGATTCAACGTGCCGACCGGGGGGTGATCCGAAAGAGCTCGCT	1447
Q	y	1474	GCGACGCTGGGACAGCTGCCGGAATTATAGTATGAGGGTTACTCCGATGGAATGCCA	1533
D	b	1448	GTGATGCGTGGGGCCCACTGCACCGACACAGGATAGGTCAACTCACTGAGTGGACGCCG	1507
Q	y	1534	CCGACCAATTCACCGGCAAAAAACAGTTCTGCAAGGCCCTCTCTGGGTCTGACAGTG	1593
D	b	1508	GCCACCAATTCACGATGCAAGAACAGTTTTCGAAAGCCCTCTTCTGGGTCTGGAAGTG	1567
Q	y	1594	TCACGACTGTGGGGACGGAAGTGAAGAGAGGGCTGCAGCTGTCTGTGGAGATTTC	1653
D	b	1568	TGAACGACTGCGGAGACAAACAGCAGACAGAGGGGTGAGTTGTCCGGGCCAAGCTTCA	1627
Q	y	1634	AGTGTTCATATGGAAAGTGTCTCCCTCAAGCAGCAGAAAGTAAATGGAAAGGAACCTGTG	1713
D	b	1628	GGTGTTTCATATGGAAAGTGTCTCTGAAAAGCCAGAGTGCATATGGAAAGACAGCTGTG	1687
Q	y	1714	GAGATGGGCTGACGAGGCTTCATGTGACAGCGCTGATGTGATCTCTTGACCAAAATA	1773
D	b	1688	GGGACGGGTCCGACGAGAGCTCTCTGCCCCAAGGTGAACGTGCTCACTGTACCAAAACA	1747
Q	y	1774	CCTACCGCTGCCAAAATGGCCCTGTGCTGACAGCAAGGCCAACCTGATGTGATGGGAAGA	1833
D	b	1748	CTTACCGCTGCTCAATGAGGCTCTGTGTTAGCAAGGCCAACCTGATGTGACGGGAAGG	1807
Q	y	1834	CGGACTGTAGCGATGGCTCCGATGAAAGAAAACGTGACTGTGGGCTGCATCTCTTACCA	1893
D	b	1808	AGGACTGTAGCGAGGCTCAGATGAGAGAGGACTGCAGCTGTGGGCTGCGGATTCACGA	1867
Q	y	1894	AAACGGCTTCGCGTGTGTGTGTGACGAAATGCGAACGAGGCGAGTGGCCCTGTGAGGTGA	1953
D	b	1868	GACAGGCTGTGTGTGTGTGTGGGGGACAGGAATGCGATGAGGGCGAGTGGCCCTGTGAGGTAA	1927
Q	y	1954	GCTTCCAGCCCTTGGGGCCAGGGCCACTGTGTGTGGGGGCTGGGCTATCTCTCTACTGGC	2013
D	b	1928	GCTGTATCTCTGGGCCACAGGGCCAAATCTGCGGTGTTCTCTCATCTCTCCCACTGGC	1987
Q	y	2014	TGGTCTCTCAGCTCAATGTGCTTTCAAGATGACAAAATTTCAAGTACTCAGACTACAGA	2073
D	b	1988	TGGTCTCTCCGACACTGTTACATGATGACAGAGGATTCAGGTACTCAGACCCCAAGC	2047
Q	y	2074	TGTGACGCGCTTCTGTGGTCTGTGTGACCAAGACAAAGCCAGTGGCTCTGGGGGTGCAGG	2133
D	b	2048	AGTGAACGCGCTTCTGTGGGCTTCCACACACACAGGACGAGGCGGCTGTGGGGTGCAGG	2107
Q	y	2134	AGCTGAAGCTCAACGATCATGACCCACCTTCTCTCAATGATTTCACTTCGACTATG	2193
D	b	2108	AGGCGAGGCTCAAGCGCATATCTCCACCCCTTCTCAATGATTTCACTTCGACTATG	2167
Q	y	2194	ACATGCGCTTGTGAGCTGAGCTGAGAGAAAGTGGTGAAGTACAGCACCGTGTGTGGCCCATCT	2253
D	b	2168	ACATGCGGCTGTGAGCTGAGAGTGAAGAAACGGGACAGGTACAGTTCATGTTGGGCCCATCT	2227
Q	y	2254	GCTGTCTGATGTCTACCATGATCTTCCCTGTGGCAGAGGCCATGTGGGTCAACAGCTGGG	2313
D	b	2228	GCTGTCCGACGCTTCCCATGTCTTCTCTGTCCGGAAGGCCATGTGGGTCAACAGCTGGG	2287
Q	y	2314	GGCACACAAAGAGGAGTACCGGACGCTGATCTTGCAGAAAGGTGAGTCCGTGTCA	2373
D	b	2288	GACACACCAATGTAGAGGACTGTGGGGGTGATCTTGCAAAAGGTTGAGATCCGCTCA	2347
Q	y	2374	TCAACGACCACTGTGAGAGCTCATGCGGACGAGATCACCCCAGGAATGATGTGTG	2433

Db 2348 TCAACGACACCTGCGAGACCTCTGCGGACGATACGCGCGCATGATGCG 2407  
 QY 2434 TGGTTTCTCAGTGGGGGTGTGAGATCTCTGCGGAGGTGATCTGTGCGCTTGTGCA 2493  
 Db 2408 TGGCTTCTCAGGCGGGGTGTGAGATCTCTGCGGAGGTGATCTGTGCGCTTGTGCA 2467  
 QY 2494 GCGCGGAGAAAGATGGGCGAATGTTCCAGGCTGTGTGTGAGCTGGGGTGAAGGCTGCG 2553  
 Db 2468 GCGTGAAGGCGGATGGGCGGATCTTCCAGGCGGGTGTGTGAGCTGGGGTGAAGGCTGCG 2527  
 QY 2554 CTGAGAGGAAAGGAGGCGGTGTGAGCAAGGCTCTGCTGTGAGTTCGGGACTGTGATCAAG 2613  
 Db 2528 CTGAGAGGAAAGGAGGCGGTGTGAGCAAGGCTCTGCTGTGAGTTCGGGACTGTGATCAAG 2587  
 QY 2614 AGCAGCTGGGATATGACGATGACACACA 2645  
 Db 2588 AGAACACTGGGATATGAGGCGCGGACACCA 2619

RESULT 4  
 AAX87815  
 ID AAX87815 standard; cDNA; 3147 BP.  
 AC AAX87815;  
 DT 09-NOV-1999 (first entry)  
 DE Tumour antigen derived gene-15 (TADG-15) cDNA.  
 KW Tumour antigen derived gene-15; TADG-15; serine protease; human;  
 KW breast cancer; ovary cancer; carcinoma; diagnosis; ss.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FT CDS 23..2590  
 FT /\*tag= a  
 PN MO9942120-A1.  
 PD 26-AUG-1999.  
 PF 18-FEB-1999; 99WO-US003436.  
 PR 20-FEB-1998; 98US-00027337.  
 PA (UTAR-) UNIV ARKANSAS.  
 PI O'Brien TJ, Tamimoto H;  
 DR MPI: 1999-527418/44.  
 DR P-PSDB; AAY06671.  
 PS A new extracellular serine protease for diagnosis of neoplastic disease.  
 XX Claim 2; Fig 9; 71pp; English.  
 CC This is the nucleotide sequence of an isolated cDNA that codes for an  
 CC extracellular serine protease, termed tumour antigen derived gene-15  
 CC protein (see AAY06671), that is overexpressed in breast and ovarian  
 CC carcinomas. The TADG-15 gene can be used as a diagnostic and therapeutic  
 CC target in ovarian carcinoma and other carcinomas including breast,  
 CC prostate, lung and colon. The TADG-15 cDNA was isolated from ovarian  
 CC carcinoma by PCR using primers directed to conserved areas of the serine  
 CC protease family. The invention also provides: a vector that is capable of  
 CC expressing DNA encoding TADG-15 protein; host cells selected from  
 CC bacterial cells (especially Escherichia coli), mammalian cells, plant  
 CC cells and insect cells; and a method of detecting expression of TADG-15  
 CC protein using a hybridisation probe  
 SQ Sequence 3147 BP; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other;

Best Local Similarity 81.2%; Pred. No. 0;  
 Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;  
 QY 45 GATGGAGCGCCAAACCAATGGGTAGCAATCGGGGCGCCGACGCGGAGGGGCTTCTAG 104  
 Db 5 GAGCGGCTTCGGGGTACCAATGGGGAGCGATCGGGCCCGGACGAGGGGCGCCGAG 64  
 QY 105 GACTTCGGCGCGGAGCTCAAGTACATCCCGGCTAGAGAACTGATGCTTTGAGAG 164  
 Db 65 GACTTCGGCGCGGAGCTCAAGTACATCCCGGAGAGAGAAATGATGCTTTGAGAG 124  
 QY 165 GGTGTGAGTTCCTGCTCGGAACTTCGCAAGAAAGTGAAGAGGAGGCGCCAGGCG 224  
 Db 125 GGGTGAAGTTCCTGCTCGGAACTTCGCAAGAAAGTGAAGAGGAGGAGGCGCCAG 184  
 QY 225 TGGGTGAGTTCCTGCTCGGAACTTCGCAAGAAAGTGAAGAGGAGGAGGCGCCAG 284  
 Db 185 TGGGTGAGTTCCTGCTCGGAACTTCGCAAGAAAGTGAAGAGGAGGAGGCGCCAG 244  
 QY 285 CTGAGTGGCACTTCATATCGAATGCGGGGTTCAAAAGTCTTCAATGGCCATTC 344  
 Db 245 CTGAGTGGCACTTCATATCGAATGCGGGGTTCAAAAGTCTTCAATGGCCATTC 304  
 QY 345 AGGATCAAAATGAGATCTTTCTGATGCGTATGAGAACTTCACTCAAGATTTATC 404  
 Db 305 AGGATCAAAATGAGATCTTTCTGATGCGTATGAGAACTTCACTCAAGATTTATC 364  
 QY 405 AGCTGGCGGAGCGAGTGAAGAGGCGGTGAAGCTGCTGTACATGAATCTCTGCTG 464  
 Db 365 AGCTGGCGGAGCGAGTGAAGAGGCGGTGAAGCTGCTGTACATGAATCTCTGCTG 424  
 QY 465 GGTCTCTTCAACAGAAAGTGGCTGTAACTGCTTCACTGAGGCGAGTGTCAATGCT 524  
 Db 425 GGTCTCTTCAACAGAAAGTGGCTGTAACTGCTTCACTGAGGCGAGTGTCAATGCT 484  
 QY 525 TACTGTGAGATTCAGATTCCTCCCACTGCGAGAAAGGATTGATGCGCATGCT 584  
 Db 485 TACTGTGAGATTCAGATTCCTCCCACTGCGAGAAAGGATTGATGCGCATGCT 544  
 QY 585 GTGAGCGAGTGTGAATGCTGACCGCGGAGAGCGGCACTGAATCTCTGCTGTAACA 644  
 Db 545 GTGAGCGAGTGTGAATGCTGACCGCGGAGAGCGGCACTGAATCTCTGCTGTAACA 604  
 QY 645 TCTGTGAGTGGCTTCCCATTTGACCCCAAGATGCTGACAGAGACTCAGACAAAGCTG 704  
 Db 605 TCAGTGTGAGTGGCTTCCCATTTGACCCCAAGATGCTGACAGAGACTCAGACAAAGCTG 664  
 QY 705 AGTTTGGCTGATGCGATGCGATGCGAGAGTGAAGAGCTTCACTTACCTTGGCTTCC 764  
 Db 665 AGTTTGGCTGATGCGATGCGAGAGTGAAGAGCTTCACTTACCTTGGCTTCC 724  
 QY 765 AACAGTCCCTTCCCGGCGATGCGCGGCGGAGTGGTCTGCGGGGGGAGAGCGGCACT 824  
 Db 725 GACAGCCCTTCCCGGCGATGCGCGGCGGAGTGGTCTGCGGGGGGAGAGCGGCACT 784  
 QY 825 GTGCTGACCTTCCCGGCGATGCGCGGCGGAGTGGTCTGCGGGGGGAGAGCGGCACT 884  
 Db 785 GTGCTGACCTTCCCGGCGATGCGCGGCGGAGTGGTCTGCGGGGGGAGAGCGGCACT 844  
 QY 885 CTGCTGACCTTCCCGGCGATGCGCGGCGGAGTGGTCTGCGGGGGGAGAGCGGCACT 944  
 Db 845 CTGCTGACCTTCCCGGCGATGCGCGGCGGAGTGGTCTGCGGGGGGAGAGCGGCACT 904  
 QY 945 GAGACCTTCCCGGCGATGCGCGGCGGAGTGGTCTGCGGGGGGAGAGCGGCACT 1004  
 Db 905 GAGACCTTCCCGGCGATGCGCGGCGGAGTGGTCTGCGGGGGGAGAGCGGCACT 964  
 QY 1005 AGCTGATTAACCAATGATGAGCGGAGAGTGGCTTGAAGGCACTTTCAGAGT 1064  
 Db 965 ACCTGATTAACCAATGATGAGCGGAGAGTGGCTTGAAGGCACTTTCAGAGT 1024  
 QY 1065 CCCAAGATGAGAGTGGCTTGAAGGCACTTTCAGAGT 1124

Query Match 60.6%; Score 1883.2; DB 2; Length 3147;



Db 1025 CTTAGGATGAGCAGCTGTGAGGCGCTTACGTAAGCCCAAGGAGCAATTCAACAGCCCC 1084  
 QY 1125 TACTATCCAGGCGCACTACCCGCCCAACATCACTGCACATGAAATATCAAGGTGCCAAC 1184  
 Db 1085 TACTACCCAGGCGCACTACCCCAACATCACTGCACATGAAATATGAGGTGCCAAC 1144  
 QY 1185 AACCGGAAAGTGAAGGTGGGCTTCAAACTCTTATCTGGTGAACCCCAACGTAACAGT 1244  
 Db 1145 AACCAACATGTGAAGGTGAATTTCTTACTCTGAGGCCCGGCTGCTGCG 1204  
 QY 1245 GGTCTCTGACCAAGGACTATGTGAGATCAACGAGGAGAAATCTCCGCTGAAGCTCG 1304  
 Db 1205 GGCACCTGCCCCAAGAGACTACGTGAGATCAATGGGAGAAATATCGCGAGAGAGCTCC 1264  
 QY 1305 CAGTTGTGTGAGCAGCAACAGAGACAGATTACAGTCCACTTCCATTCTGATCTCG 1364  
 Db 1265 CAGTTGTGTGAGCAGCAACAGATCAAGATTCAGTTGCTTCACTCACTGATCTGCTC 1324  
 QY 1365 TACACGGAACACGAGGTTCTAGCTGAGTACCTCTCTAGACTCCAAACGACCGTGCCA 1424  
 Db 1325 TACACGGAACACGAGGTTCTAGCTGAGTACCTCTCTAGACTCCAAACGACCGTGCCA 1384  
 QY 1425 GGGATGTTCAATGTGCAAGACTGAGACGCTGCATCCGAAAGGAATCGCGCTGCGA 1484  
 Db 1385 GGGCAAGTTCAAGTCCGCAACGAGGCGGTGATCCGGAAGAGCTGCGCTGATGAGCTGG 1444  
 QY 1485 GCAGACTGCCGGAATTAATGATGAGCTTACTGCCGATGCAATGCAACCCACCACTTC 1544  
 Db 1445 GCCGACTGCACCAACACAGCTGAGTCACTCACTGAGTTCGACGCGGCGCACTTC 1504  
 QY 1545 ACCGTCGAAACCAAGTTCTGCAAGCCCTCTCTCTGGTCTGTGACAGTGTCAAAGCTGT 1604  
 Db 1505 ACCGTCGAAACCAAGTTCTGCAAGCCCTCTCTCTGGTCTGTGACAGTGTCAAAGCTGT 1564  
 QY 1605 GGGGACGGAAGTGAAGAGAGGCTGACGCTGTCTCTGGGAGTTTCAAGTGTCCAAT 1664  
 Db 1565 GGAAGCAACAGCGACGAGCGAGGTGCAAGTGTGCGGCCCAAGACCTTCAGGTGTCCAAT 1624  
 QY 1665 GGGAGGTGTCTCCCTCAGAGCCAGAGTGTATGGAAGAGCACTGTGAGATGGGTCT 1724  
 Db 1625 GGGAGGTGTCTCCCTCAGAGCCAGAGTGTATGGAAGAGCACTGTGAGATGGGTCT 1684  
 QY 1725 GACGAGGCTTCATGACAGGCTGATGTGCTCTTTCACCAATATACCTACCGCTGC 1784  
 Db 1685 GACGAGGCTTCCTCCCAAGGTGAAGTGTCTCTTACCAATACCACTTACCGCTGC 1744  
 QY 1785 CAAATAGGCTCTGTCTGAGCAAGGCAACCTGATGTGTATGGAAGAGCGAAGCTGAGC 1844  
 Db 1745 CTCAATGGGCTCTGCTTGAAGCAAGGGCAACCTGATGTGTAGCGGAAGAGAGCTGAGC 1804  
 QY 1845 GATGGCTCCGATGAAATCTGTGACTGTGGGCTGCGATCTTTTACCAACAGGCTGCG 1904  
 Db 1805 GACGCTCAGATGGAAGAGCTGCACTGTGGCTGCGGTATTCAGCAGACAGGCTGCT 1864  
 QY 1905 GTGTTGGTGGCAGGAATGCGGACGAGGGGAGTGGCCCTGCAAGTGAAGCTCCACGCC 1964  
 Db 1865 GTTGTGGGGGCAAGATGCGGATGAGGGCAAGTGGCCCTGCAAGTGAAGCTCCACGCC 1924  
 QY 1965 CTGGGCGCAGGGCCATCTGCGGTGTCTCCCTCATCTCTCCCACTGGCTGTCTGCGC 2024  
 Db 1925 CTGGGCGCAGGGCCATCTGCGGTGTCTCCCTCATCTCTCCCACTGGCTGTCTGCGC 1984  
 QY 2025 GCTCATTTGCTTCAAGATGACAAAAATTTCAAGTACTCAGACTACAGATGAGTGAAGCGCC 2084  
 Db 1985 GCACACTGCTCATCTGATGACAGAGATTCAAGTACTCAGACTCAGAGCCGACAGTGAAGCGCC 2044  
 QY 2085 TTCTGGGCTCTGAGCAGAGCAAGGCAAGGCAAGTGGCTCTGGGAGTGAAGCTG 2144  
 Db 2045 TTCTGGGCTCTGAGCAGAGCAAGGCAAGGCAAGGCAAGTGGCTCTGGGAGTGAAGCTG 2104  
 QY 2145 AAAGCATATCACCACCTCTCTTCAATGATTTACCTTGAATGACTGAGTGCCTTG 2204  
 Db 2105 AAAGCATATCCTCCACCCCTTCTTCAATGACTTCACTTGAATGACTGAGTGCCTTG 2164

QY 2205 CTGAGCTGGAAGAGTGGTGAAGTACAGACCGTGTGTCGCCCATCTGCTGCTGAT 2264  
 Db 2165 CTGAGAGTGAAGAAACCGCAGAGTACACTGATGTGTGGCCCATCTGCTGCGGAGC 2224  
 QY 2265 GCTAACCATGTCTTCCCTGCTGCGAAGGCCATCTGGGTCAAGAGCTGGGGGCAACAAAA 2324  
 Db 2225 GCTTCCATGTCTTCCCTGCTGCGAAGGCCATCTGGGTCAAGAGCTGGGGGCAACACCAG 2284  
 QY 2325 GAGGAGGTAAACCGAGCGCTGATCTCTGCAAGAGGTGAATCCGTGTCAATCAACAGACC 2384  
 Db 2285 TATGAGGCACTGGCGCGCTGATCTGCAAAAGGTGAATCCGTGTCAATCAACAGACC 2344  
 QY 2385 ACCGTGAGGAACTCATGCGCGCAGCAGATCACCCCAAGATGATGTGTGGTTCCTC 2444  
 Db 2345 ACCGTGAGGAACTCATGCGCGCAGCAGATCACCCCGGATGATGTGTGGTTCCTC 2404  
 QY 2445 AGTGGAGGTGTGACTCTGCTGCAAGGTGACTCTGTGTGCCCCCTTGTTCAGCGGAGAA 2504  
 Db 2405 AGCGCGGCGTGAATCTGCTGCGAGGGTGATTCGAGGAGACCCCTGTCCAGCGTGAAGCG 2464  
 QY 2505 GATGGCGAATGTTCAAGCTGTGTGTGAGCTGGGAGTGAAGGCTGCGCTCAGAGAAC 2564  
 Db 2465 GATGGCGAATCTTCAAGCTGTGTGTGAGCTGGGAGTGAAGGCTGCGCTCAGAGAAC 2524  
 QY 2565 AACCCAGCGTGTACACAAAGGCTCCCTGTAGTTGGGACTGATCAAAAGACACTGGG 2624  
 Db 2525 AACCCAGCGTGTACACAAAGGCTCCCTGTAGTTGGGACTGATCAAAAGAACACTGGG 2584  
 QY 2625 GTATAGCAGATGACAGACAGACCGGACCAACCAACAGAGGATCCGCAATGAGCA 2684  
 Db 2585 GTATAGGAGGCGCGG--GCCACCCAAATGTGTACACTGTGGGGCCACCCATGTCCACC 2641  
 QY 2685 CCGATATACAGAGAGGAACTGACAGCAATTATGCTGTGGCTCCCGCCCAACACA 2744  
 Db 2642 CCAGTGTGACG--CTGCAAGCTGGAAGCTGGAACCGCTGACTGACACAGGCGC--CCAGA 2699  
 QY 2745 ACCCAGACTGTGAACCTGATCTTGAAGCTGAGAT 2780  
 Db 2700 ACATAGACTGTGAACCTGATCTTGAAGCTGAGAT 2735

RESULT 5  
 AAH23609/c  
 ID AAH23609 standard; RNA; 3147 BP.  
 AC AAH23609;  
 DT 03-AUG-2001 (first entry)  
 XX  
 DE Human TADG-15 antisense RNA sequence.  
 KM TADG-15; cytosolic; vaccine; ovarian tumour; cancer; human; antisense;  
 KM tumour antigen-derived gene 15; serine protease; ss.  
 OS Homo sapiens.  
 XX  
 PN W0200129056-A1.  
 XX  
 PD 26-APR-2001.  
 XX  
 PF 20-OCT-2000; 2000WO-US029095.  
 XX  
 PR 20-OCT-1999; 99US-00421213.  
 XX  
 PA (UTAR-) UNIV ARKANSAS.  
 XX  
 PI O'Brien TV, Tanimoto H;  
 XX  
 DR WPI, 2001-381031/40.  
 XX  
 PT Novel extracellular serine protease, termed tumor antigen-derived gene 15 protein overexpressed in carcinomas and DNA encoding it, for diagnosis,



XX treatment, prevention of cancer, particularly breast, ovarian cancer.  
XX  
PS Example 8; Page 102-103; 130pp; English.

CC The present invention relates to human tumour antigen-derived gene 15  
CC (TADG-15) protein and coding sequence (see AAH21601 and AB98500). TADG-  
CC 15 is an extracellular serine protease. It was found that TADG-15 is over  
CC expressed in ovarian tumours. TADG-15 protein or its fragments of 9-20  
CC residues that lack TADG-15 protease activity are useful for vaccinating  
CC an individual against TADG-15, having, suspected of having or at risk of  
CC getting cancer. Furthermore, the TADG-15 gene can be used as a diagnostic  
CC or therapeutic target in cancer. TADG-15 was cloned and expressed in the  
CC opposite orientation such that an antisense RNA molecule (the present  
CC sequence) was produced. The present sequence is useful for hybridising to  
CC the complementary RNA in a cell and thereby inhibiting translation of  
CC TADG-15 into protein

XX Sequence 3147 BP; 583 A; 958 C; 952 G; 0 T; 654 U; 0 Other;

Query Match 60.6%; Score 1883.2; DB 4; Length 3147;  
Best Local Similarity 81.2%; Pred. No. 0;  
Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;

QY	45	GATCGACCGCCAAACCATGGGTAGCATCGGGCGCGGAGGCCGAGGGGCTCTG	104
DB	3143	GAGCGGCTCGGGGATACCATGGGAGCCATCGGAGCGGAGGGGCGGAGGCGGAGG	3084
QY	105	GACTTCGGCGGGGAGCTCAAGTAACTCCCGCTAGAGACATGATGGCTTTGAGAG	164
DB	3083	GACTTCGGCGGGGAGCTCAAGTAACTCCCGGACAGAAAGTGAATGGCTTGAAGAA	3024
QY	165	GGTGGAGGTTCTCTGCTGCGGAAATGCGAAGAAAGTGAAGAGGAGGCGCCAGGCGC	224
DB	3023	GGCGTGGAGTTCTCTGCACTCAACAACTGCAAGAAAGTGAAGAGGAGGCGCGC	2964
QY	225	TGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	284
DB	2963	TGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2904
QY	285	CTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	344
DB	2903	CTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2844
QY	345	AGGATCAAAATGAGATCTTTCTGATGCGTATGAAACTCCACTCCACAGAGTTTATC	404
DB	2843	AGGATCAAAATGAGATCTTTCTGATGCGTATGAAACTCCACTCCACTGAGTTTGA	2784
QY	405	AGCCTGCGCAGCGAGTGAAGAGGCGCTGAAGCTGCTGTAATGAAGTCCCTGCTG	464
DB	2783	AGCCTGCGCAGCGAGTGAAGAGGCGCTGTAAGCTGCTGTAAGGAGTCCCATTTCTG	2724
QY	465	GATCCCTAACAAGAGTGGCTGTAACTGCTTCACTGAGGGGAGTGTACGCTTAC	524
DB	2723	GATCCCTAACAAGAGTGGCTGTAACTGCTTCACTGAGGGGAGTGTACGCTTAC	2664
QY	525	TACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	584
DB	2663	TACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2604
QY	585	GTGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	644
DB	2603	GAGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2544
QY	645	TCTGTGTGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC	704
DB	2543	TCACTGTGTGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC	2484
QY	705	AGTTTGTGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC	764
DB	2483	AGCTTGTGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCC	2424
QY	765	AACAGTCCCTAACCGGCGCATGCGCGCTGAGTGGGCTCTGCGGGGGGAGCGCGACTCT	824

DB	2423	GACAGCCCTTACCCCGCTCATGCCGCTGCGAGTGGGCGCTTGGGGGAGCGCGACTCA	2364
QY	825	GTGCTGAGCTTACCTTCCGAGGCTTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG	884
DB	2363	GTGCTGAGCTTACCTTCCGAGGCTTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG	2304
QY	885	CTGCTGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	944
DB	2303	CTGCTGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2244
QY	945	GCGACCTTTCACCTCTCTCAACCTGACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1004
DB	2243	GCGACCTTTCACCTCTCTCTCAACCTGACTTCTCTCTCTCTCTCTCTCTCTCTCTCT	2184
QY	1005	AGCGTATTAACCAATCTGACCGGCGACATCTGCTGCTTGAAGGCTTCTTCTCAAGCTG	1064
DB	2183	ACACTGATTAACCAATCTGACCGGCGACATCTGCTGCTTGAAGGCTTCTTCTCAAGCTG	2124
QY	1065	CCCAAGATGAGCAGCTGTGCGGCTTTTGAAGTACACCCAGGACATTTAGCAGCCCG	1124
DB	2123	CCTAGGATGAGCAGCTGTGAGGCGGCTTACGTTAAAGCCAGGACATTTCAACAGCCCG	2064
QY	1125	TACTATCCAGGCTCACTACCGGCTCAACATCACTGCAATGAAATATCAAGTGGCCAA	1184
DB	2063	TACTATCCAGGCTCACTACCGGCTCAACATTTGACTGCAATGAAATATGAGGTGCCAA	2004
QY	1185	AACCGGAACGTGAAGTGGCTTCAACCTCTCTATCTGCTGAGACCCCAAGTACCAATG	1244
DB	2003	AACCGGACGTGAAGTGGCTTCAACCTCTCTATCTGCTGAGACCCCAAGTACCAATG	1944
QY	1245	GGCTCTGACCAAGGACTATGATGAGATCAACGGGGAGAACTACTGCGGTGAGAGTCC	1304
DB	1943	GGCACCTGCGCCCAAGGACTATGATGAGATCAATGGGGAGAAATCTGCGGAGAGAGTCC	1884
QY	1305	CAGTTGTGTGAGGAGCAACAGCAAGATTAAGTCACTTCCTCATCTGATCACTG	1364
DB	1883	CAGTTGTGTGAGGAGCAACAGCAAGATTAAGTCACTTCCTCATCTGATCACTG	1824
QY	1365	TACACGGAACCGGGTCTTGTGAGTACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1424
DB	1823	TACACGGAACCGGGTCTTGTGAGTAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1764
QY	1425	GGGATGTCATGTGACAGTCAAGGCTGATCCGAAAGAACTGCGCTGCGAGCGCTGG	1484
DB	1763	GGGAGGTTCAAGTGTGCGGACCGGGGCTGTATCCGGAAGAGTCTGCTGTATGCTGG	1704
QY	1485	GCAAGCTCCCGGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1544
DB	1703	GCCGACTGCAACGACCAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT	1644
QY	1545	ACGTGCAAAACCAAGTCTGCAAGCCCTCTTCTGCGGTCTGTGACAGTGTCAACACTGT	1604
DB	1643	ACGTGCAAAACCAAGTCTGCAAGCCCTCTTCTGCGGTCTGTGACAGTGTCAACACTGT	1584
QY	1605	GGGAGCGGAATGACAGAGAGGCTGAGCTGTCTGCTGCGGAGTTTCAAGTGTCCAT	1664
DB	1583	GGAAGCAACAGCGACAGAGAGGCTGAGCTGTCTGCGGAGCTTCAAGTGTTCAT	1524
QY	1665	GGGAGAGTGTCTTCTCTCAAGCTCAAGATGTAATGGAAGCAACTGTGAGATGGTCT	1724
DB	1523	GGGAGAGTGTCTTCTTGAAGAGCAAGTGCATGGAAGGAGCACTGTGGGAGCGGCTTC	1464
QY	1725	GACGAGGCTTATGACAGGCTGAATGCTCTCTTGAACCAATTAATTAACCTACCGCTGC	1784
DB	1463	GACGAGGCTTCTCTCCCAAGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1404
QY	1785	CAAAATGAGCTCTGCTGAGCAAGGCAACCTGAGTGTGATGGAAGAGCGACTGAGC	1844
DB	1403	CTCAATGAGGCTCTGCTTGAAGAGGCAACCTGAGTGTGATGGAAGAGAGCTGAGC	1344
QY	1845	GATGCTCCCATGAGAAACTGTGATGCTGTGGCTGTGATCTTTTCAACAAACAGGCTGC	1904
DB	1343	GACGCTCAAGTGAAGGACTGAGCTGTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG	1284

```

QY 1905 GTGTTGTGTCGACGAGATGCGACGAGGCGGAGTGCCCTTGGCAGGTGAGCTTCCACGCC 1964
XX |||
OS |||
Db 1283 GTTGTGGGGGCGACGATGCGGATGAGGGCGAGTGCCCTTGGCAGGTGAGCTTCCATGCT 1224
QY 1965 CTGGGCGACGAGGCGCACTGTGTGGGGGCTGAGCTCATCTCTCTGACGTGGTGTCTTGCA 2024
XX |||
FT |||
Db 1223 CTGGGCGACGAGGCGCACTGTGTGGGGGCTTCCCTCATCTCTCTGACGTGGTGTCTTGCC 1164
QY 2025 GCTCATTTGCTTTCAGAGTACCAAAAATTTCAAGTACTCAGACTACAGATGTGGACGGCC 2084
XX |||
OS |||
Db 1163 GCACACTGTCTACATGATGACAGAGATTGAGTACTCAGACCCCGACGAGTGGACGGCC 1104
QY 2085 TTCTCTGGGTCTGTGACACGACGACGAGGCTCTGGGGTGTGAGAGAGCTGAGAGCTC 2144
XX |||
FT |||
Db 1103 TTCTCTGGGTCTGTGACACGACGAGGCTCTGGGGTGTGAGAGAGCTGAGAGCTC 1044
QY 2145 AAACGTATCATCACCCGACCTTCTTCAATGATTTTCACTTTCAGCTATGACATGCGCTTG 2204
XX |||
OS |||
Db 1043 AAGCGCATCATCTCCACACCCCTTCTTCAATGATTTTCACTTTCAGCTATGACATGCGCTTG 984
QY 2205 CTGGAGCTGTGAGAAAGTGTGTGAGTACAGCAGCTGTGTGCGCCCATCTGCTGCTGAT 2264
XX |||
OS |||
Db 983 CTGGAGCTGTGAGAAAGTGTGTGAGTACAGCAGCTGTGTGCGCCCATCTGCTGCGGAC 924
QY 2265 GCTACCCATCTCTTCCCTGCTGTGCGAAGGCGCATGCGGCTCACAGGCTGGGGGCGACACAAAA 2324
XX |||
OS |||
Db 923 GCTCTCCATGTCTTCTCTGCGGCGAGGCGCATGCGGCTCACAGGCTGGGGGCGACACCCAG 864
QY 2325 GAGGAGGTACCCGAGCGGTGATCTCTGACAGAGGAGTGTGATCTGTCATCAACACAGACC 2384
XX |||
OS |||
Db 863 TATGGAGAGCACTGGCGGCTGATCTCTGCAAAAGGATGATCTGCGCGTCATCAACACAGACC 804
QY 2385 ACCGTGAGGACCTCATGCTGCGGACGAGTACACCCAGAAATGATGTGTGGGTTTCTC 2444
XX |||
OS |||
Db 803 ACTGCGAGAACTCTCTGCGGACGAGTACACCCAGAAATGATGTGTGGGTTTCTC 744
QY 2445 AGTGGGGGTGTGACCTCTGCGGAGGAGTGTGATCTGATGAGGCGCCCTTGTCAAGGCGGAGAAA 2504
XX |||
OS |||
Db 743 AGGCGCGGCGTGTGACCTCTGCGGAGGAGTGTGATCTGATGAGGCGCCCTTGTCAAGGCGGAGAAA 684
QY 2505 GATGGGCGAAATGTTCCAGGCTGTGTGTGTGAGTGTGAGGCTGTGAGGCTGTGAGGAGAAC 2564
XX |||
OS |||
Db 683 GATGGGCGGATCTTCCAGGCGCGGTGTGTGTGAGTGTGAGGCTGTGAGGCTGTGAGGAGAAC 624
QY 2565 AAGCCAGGCGGTGTACACAAAGGCTCTCTGTGTGTGAGTGTGAGGCTGTGAGGCTGTGAGGAGAAC 2624
XX |||
OS |||
Db 623 AAGCCAGGCGGTGTACACAAAGGCTCTCTGTGTGTGAGTGTGAGGCTGTGAGGCTGTGAGGAGAAC 564
QY 2625 GTATAGCAGATGTGACAGACGACGACCAACACACGAGGAGTGGCGGACATGACCA 2684
XX |||
OS |||
Db 563 GTATAGGAGGCGGCGG--GCCACCCAAATGTGTACACCTGCGGAGGCGACACCATGTCCACC 507
QY 2685 CCTGGATACAGAGAGAGAACCTGACGACATTTATGTGTGGGCTTCCCCCCCCCAACACA 2744
XX |||
OS |||
Db 506 CCAAGTGTGACAG--CCTGACAGGCTGGAAGCTGGAACGCGCTGACTGACACGACGCGCC--CCAGA 449
QY 2745 ACCCAGACTGTGAAGTGTGATCTTATGAGACTCAGAGT 2780
XX |||
OS |||
Db 448 ACATACACTGTGAACCTCAATCTCCAGGGGCTCCAAAT 413

```

RESULT 6  
AAH23601  
ID AAH23601 standard; DNA; 3147 BP.

XX AAH23601;

XX 03-AUG-2001 (first entry)

XX Human TADG-15 coding sequence.

XX Human; TADG-15; cytosolic; vaccine; ovarian tumour; cancer;

```

KW tumour antigen-derived gene 15; extracellular serine protease; ds.
XX Homo sapiens.
XX Key Location/Qualifiers
XX CDS 23..2590
XX FT /*tag=a
XX FT /product="Human TADG-15"
XX
XX MO200129056-A1.
XX
XX 26-APR-2001.
XX
XX 20-OCT-2000; 2000MO-US029095.
XX
XX 20-OCT-1999; 99US-00421213.
XX
XX (UYAR-) UNIV ARKANSAS.
XX
XX O'brien TJ, Tanimoto H;
XX
XX WPI: 2001-381031/40.
XX
XX DR P-PSDB; AAB98500.
XX
XX PT Novel extracellular serine protease, termed tumor antigen-derived gene 15
XX treatment, prevention of cancer, particularly breast, ovarian cancer.
XX
XX Claim 2; Fig 2; 130pp; English.
XX
XX The present sequence is the coding sequence for human tumour antigen-
XX derived gene 15 (TADG-15) protein. TADG-15 is an extracellular serine
XX protease. It was found that TADG-15 is over-expressed in ovarian tumours.
XX TADG-15 protein or its fragments of 9-20 residues that lack TADG-15
XX protease activity are useful for vaccinating an individual against TADG-
XX 15, having, suspected of having or at risk of getting cancer.
XX Furthermore, the TADG-15 gene can be used as a diagnostic or therapeutic
XX target in cancer
XX
XX Sequence 3147 BP; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other;
XX
XX
XX Query Match 60.6%; Score 1883.2; DB 4; Length 3147;
XX Best Local Similarity 81.2%; Pred. No. 0;
XX Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;
XX
XX 45 GATGCGACCGCCAAACCATGAGTACGATCGGCGGACGAGGCGGCTCTCAG 104
XX |||
XX 5 GAGCGGCTCGGGGTACCATGCGGAGCGATCGGCGCGCAAGGCGGAGGCGCCGAAAG 64
XX |||
XX 105 GACTTCGCGCGGAGTCTCAAGTACACTCCCGGCTAAGAAATGATGTGTGAGAG 164
XX |||
XX 65 GACTTCGCGCGGAGTCTCAAGTACACTCCCGGCGAGAGAAAGTGTGTGAGAGAA 124
XX |||
XX 165 GCTGTGAGATTCTCGCGGAGAACATGCGCAAGAAAGTGTGAGAGAGGCGCCAGGCGC 224
XX |||
XX 125 GCGGTGAGATTCTCGCGGAGAACATGCGCAAGAAAGTGTGTGAGAGAGGCGCCAGGCGC 184
XX |||
XX 225 TGGGTGTGTGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 284
XX |||
XX 185 TGGGTGTGTGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 244
XX |||
XX 285 CTGGTGTGTGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 344
XX |||
XX 245 CTGGTGTGTGTGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 304
XX |||
XX 345 AGATTCACAAATGAGATCTTCTGATGCGATGAGAACTCCACTCCAGAGATTATC 404
XX |||
XX 305 AGATTCACAAATGAGATCTTCTGATGCGATGAGAACTCCACTCCAGAGATTATC 364
XX |||
XX 405 AGCTGGCCAGCCAGGTGAGAGGCGGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 464
XX |||
XX 365 AGCTGGCCAGCCAGGTGAGAGGCGGTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 424

```

QY 465 GGTCCCTACCAAGAAGTGGCTGTAACTGCTTCACTGAGGCACTGTCATCGCTTAC 524  
DB 425 GGGCCCTACCAAGAAGTGGCTGTAACTGCTTCACTGAGGCACTGTCATCGCTTAC 484  
QY 525 TACTGTGAGATTGAGCATCCGCCCACTGCGCAGAAAGAGTTGATCGCGCATGGCT 584  
DB 485 TACTGTGAGATTGAGCATCCGCCCACTGCGCAGAAAGAGTTGATCGCGCATGGCT 544  
QY 585 GTGAGCGAGTTGTAACATTGCCACCCCGAGCACTGGAATCTCTTGTGCTAAC 644  
DB 545 GAGAGCGAGTTGTAACATTGCCACCCCGAGCACTGGAATCTCTTGTGCTAAC 604  
QY 645 TCTGTGTGGCTTCCCAATTGACCCCAAGATGCTGAGAGCTGAGACCAAGCTGC 704  
DB 605 TCAGTGTGGCTTCCCAAGATGCTGAGAGCTGAGACCAAGCTGC 664  
QY 705 AGTTTGGCTGATGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 764  
DB 665 AGCTTGGCTGATGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 724  
QY 765 AACAGTCCCTACCCCGCGCATGCTGCGCTGCGCATGCTGCGCGCGCGCATGCT 824  
DB 725 GACAGCTCCCTACCCCGCGCATGCTGCGCTGCGCATGCTGCGCGCGCGCATGCT 784  
QY 825 GTGCTGAGCTGACCTTCCGAACTGTAATGCTGCTGCTGCTGCTGCTGCTGCTG 884  
DB 785 GTGCTGAGCTGACCTTCCGAACTGTAATGCTGCTGCTGCTGCTGCTGCTGCTG 844  
QY 885 GTGCTGAGCTGACCTTCCGAACTGTAATGCTGCTGCTGCTGCTGCTGCTGCTG 944  
DB 845 GTGCTGAGCTGACCTTCCGAACTGTAATGCTGCTGCTGCTGCTGCTGCTGCTG 904  
QY 945 GGCACCTTCTACCTCTCTCAACAACCTGACTTCTCTCTCTCTCTCTCTCTCTCT 1004  
DB 905 GGCACCTTCTACCTCTCTCAACAACCTGACTTCTCTCTCTCTCTCTCTCTCTCT 964  
QY 1005 ACGCTGATTAACCAATGCTGAGCGGAGATCTGCTGCTTGAAGGCACTTCTCTCA 1064  
DB 965 ACGCTGATTAACCAATGCTGAGCGGAGATCTGCTGCTTGAAGGCACTTCTCTCA 1024  
QY 1065 CCCAAGATGAGAGCTGCTGAGCGCTTCTTGAAGTGAACAACCAAGGCACTTGAAG 1124  
DB 1025 CCTAGATGAGAGCTGCTGAGCGCTTCTTGAAGTGAACAACCAAGGCACTTGAAG 1084  
QY 1125 TACTATCCAGGCTCACTACCCGCCCAATCACTGCAATGAGATATCAAGGCTGCC 1184  
DB 1085 TACTATCCAGGCTCACTACCCGCCCAATCACTGCAATGAGATATCAAGGCTGCC 1144  
QY 1185 AACCGAAGTGAAGTGGCTGCTTGAACCTTCTTGAAGTGAACAACCAAGGCACTG 1244  
DB 1145 AACCGAAGTGAAGTGGCTGCTTGAACCTTCTTGAAGTGAACAACCAAGGCACTG 1204  
QY 1245 GGCCTCTGACCAAGATCTATGAGATGCAACGAGGAGAGTACTGCGTGAAGCTGC 1304  
DB 1205 GGCCTCTGACCAAGATCTATGAGATGCAACGAGGAGAGTACTGCGTGAAGCTGC 1264  
QY 1305 CAGTTTGTGAGAGCTGAGCAAGCAAGATTAAGTCACTTCTCACTTCTGATCTCG 1364  
DB 1265 CAGTTTGTGAGAGCTGAGCAAGCAAGATTAAGTCACTTCTCACTTCTGATCTCG 1324  
QY 1365 TACACGGAACCGGCTTCTAGCTGAGTACTCTCTCACTGAGTCAAGACCTGCTG 1424  
DB 1325 TACACGGAACCGGCTTCTAGCTGAGTACTCTCTCACTGAGTCAAGACCTGCTG 1384  
QY 1425 GAGATGTTCAATGCTGAGAGCTGAGCTGAGTCAAGAGTCAAGTCAAGCTGAG 1484  
DB 1385 GAGATGTTCAATGCTGAGAGCTGAGCTGAGTCAAGAGTCAAGTCAAGCTGAG 1444  
QY 1485 GCAACCTGCGGAGTTTATGATGAGCTGTTAATGCGAGTCAATGCTCAAGCTGCT 1544  
DB 1445 GCAACCTGCGGAGTTTATGATGAGCTGTTAATGCGAGTCAATGCTCAAGCTGCT 1504  
QY 1545 ACGTGCAAAAACGAGTCTGCAAGCCCTCTTCTGAGCTGTGACAGTCAAGCTGT 1604

DB 1505 ACGTGCAAAAACGAGTCTGCAAGCCCTCTCTTCTGAGTCTGAGTCAAGCTGCT 1564  
QY 1605 GGGGACGGAAGTGAACAGAGAGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1664  
DB 1565 GGGGACGGAAGTGAACAGAGAGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1624  
QY 1665 GGGGACGGAAGTGAACAGAGAGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1724  
DB 1625 GGGGACGGAAGTGAACAGAGAGGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1684  
QY 1725 GAGAGGCTTCAATGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 1784  
DB 1685 GAGAGGCTTCAATGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 1744  
QY 1785 CAAAATGAGCTCTGCTGAGAGAGGCAACCTGAGTGTGATGAGAGAGAGCTGAG 1844  
DB 1745 CAAAATGAGCTCTGCTGAGAGAGGCAACCTGAGTGTGATGAGAGAGAGCTGAG 1804  
QY 1845 GATGCTCCGATGAGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 1904  
DB 1805 GATGCTCCGATGAGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 1864  
QY 1905 GTGCTGAGCTGAGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 1964  
DB 1865 GTGCTGAGCTGAGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 1924  
QY 1965 GTGCTGAGCTGAGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 2024  
DB 1925 GTGCTGAGCTGAGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 1984  
QY 2025 GCTCATGCTTCAAGATGAGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 2084  
DB 1985 GCTCATGCTTCAAGATGAGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 2044  
QY 2085 TTCTGAGTCTGCTGAGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 2144  
DB 2045 TTCTGAGTCTGCTGAGAGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 2104  
QY 2145 AAAGCTATCATCAACCACTCTCTCTCAATGATTTCACTTCTGATGATGATGATG 2204  
DB 2105 AAAGCTATCATCAACCACTCTCTCTCAATGATTTCACTTCTGATGATGATGATG 2164  
QY 2205 CTGAGCTGAGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTG 2264  
DB 2165 CTGAGCTGAGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTG 2224  
QY 2265 GCTTCCATGCTCTCTCTGCGGCAAGGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 2324  
DB 2225 GCTTCCATGCTCTCTCTGCGGCAAGGCTGAGAGCTGAGAGCTGAGAGCTGAGAG 2284  
QY 2325 GAGGAGGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTG 2384  
DB 2285 GAGGAGGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTG 2344  
QY 2385 AACTGTGAGAGCTCATGCGCGAGAGTCAACCCCAAGATGATGATGATGATGATG 2444  
DB 2345 AACTGTGAGAGCTCATGCGCGAGAGTCAACCCCAAGATGATGATGATGATGATG 2404  
QY 2445 AGTGGAGGCTGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAG 2504  
DB 2405 AGTGGAGGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAG 2464  
QY 2505 GATGGAGGAGATGCTTCAAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAG 2564  
DB 2465 GATGGAGGAGATGCTTCAAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAG 2524  
QY 2565 AAGCGAGGCTGATCAAGAGCTTCTGATGAGAGAGCTGAGAGAGCTGAGAGAGCTGAG 2624  
DB 2525 AAGCGAGGCTGATCAAGAGCTTCTGATGAGAGAGCTGAGAGAGCTGAGAGAGCTGAG 2584  
QY 2625 GTATAGAGATGAGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG 2684

```

Db      2585 GTATAGGGGCGGG---GCCACCAAAATGTGTACACTGGGGGCGACCCATGTCACCC 2641
QY      2685 CCTGATATACAGAGAGAACTACTGACGACATTATGCTGTGGCTTCCCCCCCCCAACACA 2744
Db      2642 CCAGTGTGACAG-CCTGCGAGGCTGGAGACTGAGACCGCTACTGACACAGCGCC-CCAGA 2699
QY      2745 ACCCGAGCTGTGACTGACATCTCTTAGACTCAGAGT 2780
Db      2700 ACATACACTGTGACTCAATCTCCAGGGCTCCAAAT 2735

RESULT 7
AADI3155
ID      AADI3155 standard; DNA; 3147 BP.
XX      AADI3155;
XX      16-OCT-2001 (first entry)
XX      Human membrane-type serine protease (MTSP) 1 protease domain DNA.
XX      Human; transmembrane serine protease; membrane-type serine protease;
XX      MTSP; protease domain; neoplastic disease; tumour; cancer; cytostatic;
XX      lung carcinoma; colon adenocarcinoma; ovarian carcinoma; gene therapy;
XX      matrilysin; ds.
XX      Homo sapiens.
XX      OS
XX      Key      Location/Qualifiers
XX      CDS      1665..2590
XX      FT      /tag= a
XX      FT      /product= "Human transmembrane serine protease (MTSP) 1
XX      FT      /note= "CDS does not include start codon"
XX      FT      /partial
XX      PN      WO200157194-A2.
XX      PD      09-AUG-2001.
XX      PF      02-FEB-2001; 2001WO-US003471.
XX      PR      03-FEB-2000; 2000US-0179982P.
XX      PR      18-FEB-2000; 2000US-0183542P.
XX      PR      22-JUN-2000; 2000US-0213124P.
XX      PR      26-JUL-2000; 2000US-0220970P.
XX      PR      08-SEP-2000; 2000US-00657986.
XX      PR      22-SEP-2000; 2000US-0234840P.
XX      PA      (CORV-) CORVAS INT INC.
XX      PI      Madison EL, Ong EO, Yeh J;
XX      DR      WPI; 2001-488877/53.
XX      DR      P-PSDB; AAE06936.
XX      PT      Novel single chain polypeptide comprising protease domain of type-II
XX      PT      membrane-type serine protease or its catalytically active portion useful
XX      PT      for treating and preventing cancer and tumor.
XX      PS      Example 6; Page 225-227; 256pp; English.
XX      CC      The invention relates to transmembrane serine proteases and their
XX      CC      corresponding nucleotides and the protease domain of a type-II membrane-
XX      CC      type serine protease (MTSP). MTSP is useful for identifying compounds
XX      CC      that modulate or inhibit its proteolytic activity and for formulating a
XX      CC      medicament for treating neoplastic disease. MTSP and its corresponding
XX      CC      nucleotides are useful in preventing or treating tumours or cancers such
XX      CC      as lung carcinoma, colon adenocarcinoma and ovarian carcinoma, in
XX      CC      diagnostics and in hybridisation assays. MTSP is useful as a diagnostic
XX      CC      marker for tumour development, growth and/or progression and as
XX      CC      immunogens to generate antibodies that specifically bind to it. MTSP DNA
XX      CC      is useful in a yeast two-hybrid system and in gene therapy. The present

```

```

CC      sequence is a DNA encoding protease domain of human MTSP1 protein (also
CC      called matrilysin)
XX      SQ      Sequence 3147 BP; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other;
XX      Query Match      60.6%; Score 1883.2; DB 4; Length 3147;
XX      Best Local Similarity 81.2%; Pred. No. 0;
XX      Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;
QY      45 GATCGAGCCGCAAAACCATGGGTAGCAATCGGGGCGGACGAGCGGGGCTCTCAG 104
Db      5 GAGCGGCTCGGGGTACATGGGAGACGATCGGGCCCGCAAGGCGGAGGGGCCGGAAG 64
QY      105 GACTTCGGCGCGGAGACTCAAGTACAATCCCGGCTAGAGAACATGATGCTTTAGAGAG 164
Db      65 GACTTCGGCGCGGAGACTCAAGTACAATCCCGGAGAGAAAGTAAATGCTTGGAGGAA 124
QY      165 GGTGTGAGATTCTGCTCGTCAACAAATGCGCAAGAAAGTGAAGAGGAGGCGCCCAAGCGC 224
Db      125 GGGGTGAGATTCTCTGCAAGTCAACAGTCAGAAAGGTGAAAGATGAGCCCGAGCGC 184
QY      225 TGGGTGTGTGGTGGAGTGTGCTGCTCACTTCTGCTCTCCCTCATGCTGAGCTTG 284
Db      185 TGGGTGTGTGGAGCTGAGCTGCTGATCGGCTCTCTTGTGCTTGGGAGATGGCTTC 244
QY      285 CTGGTGTGCACTTCATTATCGAAATGTGGGCTTCAAAAGTCTTCAATGGCCATCTG 344
Db      245 CTGGTGTGCAATTTGAGATACCGGAGCTGCTGACAGAGGTCTTCAATGCTCATG 304
QY      345 AGGATCACAAAATAGATCTTTCTGAGATGCTATGAGAACTCCACCTCCACAGATTATC 404
Db      305 AGGATCACAAAATAGATTTTGTGAGTCTTACAGAACTCCAACTCACTGATGTTGTA 364
QY      405 AGCCTGCGCAGCAGGATGAAGAGGCGCTGAGACTGCTGTACATGAAGTCCCTGTCTG 464
Db      365 AGCCTGCGCAGGATGAAGAGAGCGCTGAAGCTGTGTACAGCGGAGTCCCATTCCTG 424
QY      465 GGTCCCTACCAAGAAAGTGGCTGTATCTGCTTCAGTGAAGGCAGTGTATCGCCTAC 524
Db      425 GGGCCCTACCAAGAGGAGTGGCTGTATCGGCTTCAGCAGAGGCGACGGTATCGCCTAC 484
QY      525 TACTGTGACAGTTTCAAGATCCGCCCAACACTGGGAGAAAGGTTGATCGCGCATGCT 584
Db      485 TACTGTGACAGTTTCAAGATCCCGACAGACTGGTGAAGAGCGCGCATGATGATG 544
QY      585 GTGAGAGGAGTTTGAATCATTTGCAACCCCGAGCACTGGCACTGAATCTTGTGTCTACA 644
Db      545 GAGAGGCGGATGATCATGCTGCCCGCGGCGGCGCTCCGAAAGTCTTGTGTGATCC 604
QY      645 TCTGTGTGAGCTTCCCATTTGACCCCAAGATGCTGCAAGAGACTCGAACAAGCTGC 704
Db      605 TCAGTGTGCTTTCCCGACGAGCTCCAAAACAGTACAGAGAGCCGAGCAACAGCTGC 664
QY      705 AGTTTGGCCCTGATGCGCAGTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 764
Db      665 AGTTTGGCCCTGATGCGCAGTGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 724
QY      765 AACAGTCCCTACCGCGGAGATGCCGCTGCGCAGTGGTCTGCGGAGGAGCGCGCATCT 824
Db      725 GACAGCCCTTACCGCGGATGCCGCTGCGCAGTGGTCTGCGGAGGAGCGCGCATCT 784
QY      825 GTGCTGAGCTTCACTTCCGAGTGGTGAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 884
Db      785 GTGCTGAGCTTCACTTCCGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 844
QY      885 CTGCTGACCGGTATGATAGCTGAGCCGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 944
Db      845 CTGCTGACCGGTATGATAGCTGAGCCGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 904
QY      945 GGACACTTTCACCTCTCTTACACCTGACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1004
Db      905 GGACACTTTCACCTCTCTTACACCTGACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 964

```

QY 1005 ACGTGATTAACCACTACTGACCGGCGACATCTGGCTTTGAGGCCACTTCTTCCAGCTG 1064  
 DB 965 AACTGTATTAACCACTACTGAGCGGCGGACATCCGGCTTTGAGGCGACCTTCTTCCAGCTG 1024  
 QY 1065 CCCAAGATGAGCAGCTGTGGCGGCTTTTGTGATGACACCCAGAGGACATTTAGAGCGCC 1124  
 DB 1025 CTTAGAGATGAGCAGCTGTGGAGCGGCTTTAGTAAAGCCAGGGGACATTTCAAGCGCC 1084  
 QY 1125 TACTATCCAGGCGCACTACCGGCCAACATCACTGACATGGAATTCAGAGTGCCCAAC 1184  
 DB 1085 TACTATCCAGGCGCACTACCGGCCAACATCACTGACATGGAATTCAGAGTGCCCAAC 1144  
 QY 1185 AACCGGAAGTGAAGGTGGCTTCAAACTCTTATCTGTGTGAGCCCAAGTACAGTG 1244  
 DB 1145 AACCAAGATGTGAAGGTGAGCTTCAAACTCTTATCTGTGTGAGCCCAAGTACAGTG 1204  
 QY 1245 GGTCTCTGACCAAGGACTATGTGAGATCAACGGGGAGAAATGACTGCGGTGAGAGTCC 1304  
 DB 1205 GGCACCTGCGCCCAAGGACTATGTGAGATCAATGGGGAGAAATGACTGCGGAGAGAGTCC 1264  
 QY 1305 CAGTTTGTGTGACAGCAACAGCAAGATTAACAGTCCACTTCCATTTGTATCACTG 1364  
 DB 1265 CAGTTTGTGTGACAGCAACAGCAAGATTAACAGTCCACTTCCATTTGTATCACTG 1324  
 QY 1365 TACACGGAACACCGGGTTCTAGCTGAGTACCTCTCTAGACCTCCACAGCCGTCCTCA 1424  
 DB 1325 TACACGGAACACCGGGTTCTAGCTGAGTACCTCTCTAGACCTCCACAGTCCATCCCG 1384  
 QY 1425 GGGATTTGATGTGACAGACTGAGCAGTGCATCCGAAGAACTGCGCTGCGACGCTGG 1484  
 DB 1385 GGGCAGTTTACGTCGCGGACGCGGGCGGTGTATCCGGAAGAGTGGCTGTGTATGCTGG 1444  
 QY 1485 GCAAGACTGCGCGGATTTATGTATGAGCGTTTACTGCGATGCAATGCCACCCACCACTTC 1544  
 DB 1445 GCGCAGTGCAGCGACCAACAGCATGAGTCAACTGCAATTGCGACGCGCGCCACCACTTC 1504  
 QY 1545 ACGTCAAAAACCAAGTCTGCAAGCCCTCTTCTGGTCTGTGACAGTGTCAACGACTGT 1604  
 DB 1505 ACGTCAAAAACCAAGTCTGCAAGCCCTCTTCTGGTCTGTGACAGTGTCAACGACTGT 1564  
 QY 1605 GGGGACGGAAGTGAAGAGAGGCGTGCAGCTGTCTGCGTGGAGTTTCAAGTGTCCAAAT 1664  
 DB 1565 GGAGCAACACGCGAGAGAGGCGGTGCAGTGTCTGCGGCCCAACCTTCAAGTGTTCAAAT 1624  
 QY 1665 GGGAACTGTCTCTCTCAGAGCCAGAAAGTGTAAATGGGAAGACACTGTGAGATGGTCT 1724  
 DB 1625 GGGAACTGTCTCTCAGAAAGCCAGAGTGCATGGGAAGACGACTGTGGAGACCGGCTCC 1684  
 QY 1725 GACGAGGCTTCAATGTGACAGCGTGAATGCGTCTCTGACCAAAATATACCTACCGCTGC 1784  
 DB 1685 GACGAGGCTCTCTGCCCCAAGGTGAACGTGTCACTTGTATCCAAACACACTACCGCTGC 1744  
 QY 1785 CAAAATGAGCTCTGTCTGAGCAAGGGCAACCTGAGTGTGATGGGAAGAGGACTGTAGC 1844  
 DB 1745 CTCAATGGGCTCTGTCTGAGCAAGGGCAACCTGAGTGTGAGGGAAGAGGACTGTAGC 1804  
 QY 1845 GATGCTCTCCGATGAGAAAACTGTGACTGTGGCTGCGATCTTTTACCAACAGGCTGC 1904  
 DB 1805 GACGCTCTCAGATGAGAAAGACTGCGACTGTGGCTGCGATCTTTTACCAACAGGCTGC 1864  
 QY 1905 GTGGTGTGTGAGCAAGAGCGGACAGAGGCGGAGTGGCTCTGCGAGTGTAGCTCCAGCGC 1964  
 DB 1865 GTTGTGTGGGGGACGGAATGGCGATGAGGCGGAGTGGCTCTGCGAGTGTAGCTCTGATGCT 1924  
 QY 1965 CTGGGCGCAAGGCGCACTTGTGTGGGCGCTGCTCATCTCTCTGACTGGCTGGTCTGCA 2024  
 DB 1925 CTGGGCGCAAGGCGCACTTGTGGGTGCTTCCCTCATCTCTCTCAACTGGCTGGTCTCTGCC 1984  
 QY 2025 GCTCATTTGCTTTCAAGATGACAAAAATTTCAAGTCTCAGACTACAGATGTGACGCGCC 2084  
 DB 1985 GCACACTCTACTATGATGACAGAGATTTCAAGTACTCAGACCCCAAGCAGTGTGACGCGCC 2044  
 QY 2085 TTCTGGGTGTGTGTGACCAAGAGCAAGGCGAGTGTCTGTGGGTGTGAGGACTGAAGCTC 2144

DB 2045 TTCTGGGTGTGTGTGACCAAGAGAGGCGGAGCGCCCTGGGTGTGAGAGCGCAGGCTC 2104  
 QY 2145 AAACGTATGATCAACCAACCTTCTCTCAATGATTTTCACTTGTGACATATGACATGCTG 2204  
 DB 2105 AAGCGATCATCTCCACCCCTTCTTCAATGATTTTCACTTGTGACATATGACATGCTG 2164  
 QY 2205 CTGGAAGTGAAGATGCTGTGAGTACAGACCGGTGTGGCGCCCATCTGCGCTGGCTGAT 2264  
 DB 2165 CTGGAAGTGAAGAACCGGACAGTACAGTCTCATGATGTGGCGCCCATCTGCGCTGG 2224  
 QY 2265 GCTACCATGTCTTCTCTGTGGCAAGCCATCTGGGTGACAGAGTGGGGGCAACAAAA 2324  
 DB 2225 GCTCCCATGTCTTCTCTGTGGCAAGCCATCTGGGTGACAGAGTGGGGGCAACACCAG 2284  
 QY 2325 GAGGAGATGACGGAGCGTGTATCTCTGCAAGAGGTGAGATCTGCTGTATCAACAGACC 2384  
 DB 2285 TATGAGGCACTGGCGCGCTGTATCTGCAAAAGGTGAGATCTGCGCTCATCAACAGACC 2344  
 QY 2385 ACCTGTGAGGACCTCATGCGCGAGAGATCAACCCCAAGATGATGTGTGGTCTCTC 2444  
 DB 2345 ACCTGTGAGAACCTCTCTGCGCGAGAGATCAACCCCGCATGATGTGTGGTCTCTC 2404  
 QY 2445 AGTGGGATGTGACCTCTGCGAGGTGACTGTGTGAGCGCCCTTGTCAAGCGCGAGAAA 2504  
 DB 2405 AGCGGCGGCTGAGCTCTGCGAGGAGTGTATCCGGGGGAGCCCTGTCCAGCGTGAAGGG 2464  
 QY 2505 GATGGGCGAATTTCCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2564  
 DB 2465 GATGGGCGAATTTCCAGGCGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2524  
 QY 2565 AAGCGAGCGTGTACAAAGGCTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2624  
 DB 2525 AAGCGAGCGTGTACAAAGGCTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2584  
 QY 2625 GTATGACAGATGACAGACGCGACCAAAACCCACAGGATGCCCGCATGTGACAG 2684  
 DB 2585 GTATGAGGCGCGGG---GCCACCAAAATGTATGACCTGCGGGGCGACCATGTGTAC 2641  
 QY 2685 CCTGATTCAGAGAGAGAAACATGACGACATTTATGTGTGTGTGTGTGTGTGTGTGTGT 2744  
 DB 2642 CCAAGTGTGAGC-CCTGAGAGGTGAGACTGTGACCGCTGTGACACAGCGGCC--CCAG 2699  
 QY 2745 ACCGAGCTGTGAACTGCATCTTGTGACTCAGAGT 2780  
 DB 2700 ACATACCTGTGACTCAATCTCCAGGGCTCCAAAT 2735  
 RESULT 8  
 AAD13113  
 ID AAD13113 standard; DNA; 3147 BP.  
 XX  
 AC AAD13113;  
 XX  
 DT 16-OCT-2001 (first entry)  
 XX  
 DE Human membrane-type serine protease (MTSP) 1 DNA.  
 XX  
 DE Human; transmembrane serine protease; membrane-type serine protease;  
 KM MTSP; protease domain; neoplastic disease; tumour; cancer; cytostatic;  
 KM lung carcinoma; colon adenocarcinoma; ovarian carcinoma; gene therapy;  
 KM matrix; ds.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 23..2590  
 FT /\*tag= a  
 FT /product= "Human transmembrane serine protease (MTSP) 1"  
 XX  
 PN W0200157194-A2.  
 XX  
 PD 09-AUG-2001.

XX 02-FEB-2001; 2001WO-US003471.  
 PF 03-FEB-2000; 2000US-0179982P.  
 XX 18-FEB-2000; 2000US-0183542P.  
 XX PR 22-JUN-2000; 2000US-0213124P.  
 PR 26-JUL-2000; 2000US-0220970P.  
 PR 08-SEP-2000; 2000US-00657986.  
 PR 22-SEP-2000; 2000US-0234840P.  
 XX (CORV-) CORVAS INT INC.  
 PA  
 PI Madison EL, Ong EO, Yeh J;  
 DR WPI; 2001-488877/53.  
 DR P-PSDB; AAE06930.  
 XX  
 PT Novel single chain polypeptide comprising protease domain of type-II  
 PT membrane-type serine protease or its catalytically active portion useful  
 PT for treating and preventing cancer and tumor.  
 PS  
 PS Claim 14; Page 191-195; 2566p; English.  
 XX  
 CC The invention relates to transmembrane serine proteases and their  
 CC corresponding nucleotides and the protease domain of a type-II membrane-  
 CC type serine protease (MTSP). MTSP is useful for identifying compounds  
 CC that modulate or inhibit its proteolytic activity and for formulating a  
 CC medicament for treating neoplastic disease. MTSP and its corresponding  
 CC nucleotides are useful in preventing or treating tumours or cancers such  
 CC as lung carcinoma, colon adenocarcinoma and ovarian carcinoma, in  
 CC diagnostics and in hybridisation assays. MTSP is useful as a diagnostic  
 CC marker for tumour development, growth and/or progression and as  
 CC immunogens to generate antibodies that specifically bind to it. MTSP DNA  
 CC is useful in a yeast two-hybrid system and in gene therapy. The present  
 CC sequence is a DNA encoding human MTSP1 protein (also called matriptase)  
 XX  
 XX Sequence 3147 BP; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other;  
 Query Match 60.6%; Score 1883.2; DB 4; Length 3147;  
 Best Local Similarity 81.2%; Pred. No. 0;  
 Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;

QY 525 TACTGTCAGAGTTCAGCATCCGCCACACCTGGCAGAAAGGTTGATCGGCCATGGCT 584  
 |||||  
 DB 485 TACTGTCAGAGTTCAGCATCCGCCAGACCTGGTGGAGAGGCCAGCGCTCATGGCC 544  
 QY 585 GTGAGACGAGTTGTAACTATGCCACCCGAGACGGGCACTGAATATCTTCGTGTAACA 644  
 |||||  
 DB 545 GAGGACGGCGTAGTCATGCTGGCCCCCGCGCGCGCTCCCTGTAAGTTCCTTGTGTACCC 604  
 QY 645 TCTGTGTGGCTTCCCACTTGAACCCAGATGCTGAGAGACTAGAGCAACAGCTGC 704  
 |||||  
 DB 605 TCAGTGTGACTTCCCAACGAGCTCAAAAACGTACAGAGACCAGAGCAACAGCTGC 664  
 QY 705 AGTTTGCCTGATGATCCCATGATGATGAGCAGTGCACGCTTCACTACCCCTGGCTTCCC 764  
 |||||  
 DB 665 AGCTTTGGCTGTGACGCCCGCGGCTGTGAGCTGATGCGCTTCAACAGCGCCGCTTCCCT 724  
 QY 765 AACAGTCCCTACCCGGCGCATGCCGCTGCTCATGGGCTCTGGGGGGGAGCGCGCATCT 824  
 |||||  
 DB 725 GACAGCCCTTACCCCGCTCATGCCCCTGTCAGATGGGCCCTGGGGGGGAGCGCGCATCA 784  
 QY 825 GTGCTGAGCCTCACCTTCCGAAGCTTGTGATGCTGCTCCCTGTGATGAGCATGGCAGTGC 884  
 |||||  
 DB 785 GTGCTGAGCCTCACCTTCCGAAGCTTGTGATGCTGCTCCCTGTGATGAGCATGGCAGTGC 844  
 QY 885 CTGGTCAACGCTGTATGATAGCTGAGCCCATGAGAACCCACGCTGTGCTGGCTGTGT 944  
 |||||  
 DB 845 CTGGTCAACGCTGTATGATAGCTGAGCCCATGAGAACCCACGCTGTGCTGGCTGTGT 904  
 QY 945 GGCACCTTCTGACCCCTCTTCAACCTGACCTTCTCTCTCCAGAAAGTCTTCTTCTGTC 1004  
 |||||  
 DB 905 GGCACCTTCTGACCCCTCTTCAACCTGACCTTCTCTCTCCAGAAAGTCTTCTTCTGTC 964  
 QY 1005 AGCGTATTAACAATCTGACCGGAGCATCTGAGCTTGTGAGGCACTTCTTCTCAAGCTG 1064  
 |||||  
 DB 965 AACCTGATTAACAATCTGACCGGAGCATCTGAGCTTGTGAGGCACTTCTTCTCAAGCTG 1024  
 QY 1065 CCCAAGATGACAGCTGTGGCGGCTTTTGTGATGACACCCAGAGGACATTTAGAGCCCC 1124  
 |||||  
 DB 1025 CTAAGATGAGAGCTGTGGAGCGCGCTTACGTAAACCCAGAGGAGCATTTCAACAGCCCC 1084  
 QY 1125 TACTATCAGGCCCACTACCCGCCCAACATCACTGACATGGAATATCAAGAGTCCCAAC 1184  
 |||||  
 DB 1085 TACTATCAGGCCCACTACCCGCCCAACATGACATGGAATATCAAGAGTCCCAAC 1144  
 QY 1185 AACCGAAACGTGAAGGTGCGCTTCAAACTCTTATCTGTGTGACCCCAACGTACAGTG 1244  
 |||||  
 DB 1145 AACCGAAACGTGAAGGTGAGCTTCAAACTCTTATCTGTGTGACCCCAACGTACAGTG 1204  
 QY 1245 GGCCTCTGACCAAGAGCTATGTGAGATCAACGGGAGAAAGTACTGCGGTGAGAGTCC 1304  
 |||||  
 DB 1205 GGCACCTGCCCAAGGACTACGTGAGATCAATGGGAGAAATTACTCGGAGAGAGTCC 1264  
 QY 1305 CAGTTTGTGTGAGACCAACAGCAGCAAGATTAAGTTCACCTTCCATTGATCACTCG 1364  
 |||||  
 DB 1265 CAGTTTGTGTGAGACCAACAGCAGCAAGATTAAGTTCACCTTCCATTGATCACTCG 1324  
 QY 1365 TACACGACACCGGGTCTTAGTGAATACCTCTCTTACGATCCCAAGACCCGTCGCCA 1424  
 |||||  
 DB 1325 TACACGACACCGGGTCTTAGTGAATACCTCTCTTACGATCCCAAGACCCGTCGCCA 1384  
 QY 1425 GGGATGTTTATGTGCAAGACTGACGCTGATCCGAAAAGAACTGCGTGGAGAGGCTGG 1484  
 |||||  
 DB 1385 GGGAGTTCACGTCGCCACGCGGCGGTGTATCCGGAAAGGAGTGCCTGTGATGGCTGG 1444  
 QY 1485 GCAAGTCCCGGATTAATAGTGAAGAGGTACTGCGGATGCAATGACCAACCAAGATTTC 1544  
 |||||  
 DB 1445 GCGCATGACCAACGACCAAGGATGAGTCAACTGCACTTCGAGGCCGCGACCAAGATTTC 1504  
 QY 1545 ACGTGCAAAAACCAATTTCTGCAAGCCCTCTTCTGGGTCTGTGACAGTGTCAACAGCTGT 1604  
 |||||  
 DB 1505 ACGTGCAAAAACCAATTTCTGCAAGCCCTCTTCTGGGTCTGTGACAGTGTGAAACAGCTGC 1564

QY	1605	GGGAGCGAAGTGAACAGAGAGGGGCTGCACTGTCTCTGGGAGTTTCAAGTGTCCAA	1664
Db	1565	GGAGACACAGGACAGAGAGGAGGTGCAATTTCTCGGCGCCAGACCTTAGAGTGTCCAA	1624
QY	1665	GGGAAGTGTCTCCCTCAGAGCCAGAAAGTGTATGGGAAAGAACATGTGGAGATGGGCT	1724
Db	1625	GGGAAGTGTCTCCGAAAAGCCAGCATGTGCATTTGGGAAAGAACATGTGGGAGAGGCTC	1684
QY	1725	GACGAGGCTTCATGTGACAGCCGTGAATGTCTCTTGCACCAATATATACCACTCGCTGC	1784
Db	1685	GACGAGGCTCTCTGCGCCCAAGGTGAAACGTCTGCTTGTACCAACACACCTTACCGCTGC	1744
QY	1785	CAAAATGCGCTTGTGTGTGACAGGCGAACCTTGAGTGTGATGGGAAACGGACTGTAGC	1844
Db	1745	CTCAATGGGCTCTGTCTTGAAGCAAGGCGAACCTTGAGTGTGACGGGAAAGAGACTGTAGC	1804
QY	1845	GATGCTCCGATGAGAAAACTGTGACTGTGGGCTGCATCTCTTACCAACAGGCTGCG	1904
Db	1805	GACGCTCAGATGAGAAAGGACTGCACTGTGGGCTGCAGTTCACAGACAGGCTGCT	1864
QY	1905	GTGTTGTGTGGCAGCAATGCGGACAGAGGCGCAGTGGCCCTGGACAGGTGAGCTCCAGCGC	1964
Db	1865	GTGTTGTGGGGGACGAGATGCGGATGAGGGCCAGTGGCCCTGGACAGTGAAGCTTCATGCT	1924
QY	1965	CTGGGCGCAGGGCCACTTGTGTGGGCGCTGCTCATCTCTCTGACTGGCTGTGTGCA	2024
Db	1925	CTGGGCGCAGGGCCACATCTGGGCTGTCTCTCTCATCTCTCCCACTGGCTGTGTGCGC	1984
QY	2025	GCTCATTGCTTTACAGATGACAAAAATTCAAGTACTCAGCTACAGATGTGACGCGC	2084
Db	1985	GCACTCTTACATGATGACAGAGATTCAAGTACTCAGCTACAGCCCAAGCTGACGCGC	2044
QY	2085	TTCTGTGGTCTGTGGACCAAGACAAGGCGCAGTGGCTGTGGGGTGCAGAGCTCAAGCTC	2144
Db	2045	TTCTGTGGCTTGCAGACAGACCAAGCCAGGCGCGCCCTGGGGTGTGCAGAGCGAGGCTC	2104
QY	2145	AAAGCTATCATCAACCACCTTCTCTTCATGATTTCACTTTCAGCTATGACATGACCTTG	2204
Db	2105	AAAGGCATCATCTCCACCCCTTCTTCATGACTTTCACCTTCGACTATGACATGGCTGT	2164
QY	2205	CTGAGGCTGAGGAAGTGTGGTGAATTAACGACCGTGTGGGCGCCCATCTGTGCTGTAT	2264
Db	2165	CTGAGGCTGAGGAACCGGCAAGTATCAAGCTCAATGATGGGCGCCCATCTGTGCGGAC	2224
QY	2265	GCTACCCATGTCTTCCCTGTGTGAGGAGCCATCTGGGTCACAGGCTGGGGGCAACAAAA	2324
Db	2225	GCTGCCATGTCTTCCCTGCGCGGCAAGGCCATCTGGGTCACAGGCTGGGGAACACCCAG	2284
QY	2325	GAGGAGGTACCGAGCGGTGTATCTTGCAGAAAGGTGAGATCCGTGTCAATCAACAGACC	2384
Db	2285	TATGAGGCACTGTGCGGCTGTATCTCTCAAAAAGGTGAGATCCGCGTCAATCAACAGACC	2344
QY	2385	ACCTGTAGGACCTCAATGCGGCAACAAATCAACCCCAAGATGATGTGTGTGGTTCTCTC	2444
Db	2345	ACCTGTAGGACCTCTCTGCGGCAACAAATCAACCGCGGCAATGATGTGTCGTTGCTCTCTC	2404
QY	2445	AGTGGGGGTGTGACTCTGTGCACAGGCTGACTCTGTGTGGCCCTTGTCAACGCGGAGAAA	2504
Db	2405	AGCGGCGGCGTGAAGCTCTGTGCAGGCTGTATTCGGGGGAACCCCTGTCAACGCGGAGCG	2464
QY	2505	GATGGCGAATGTTCCAGGCTGTGTGTGAGCTGGGGTGAAGGCTGCTCAGAGAAAC	2564
Db	2465	GATGGCGAATCTTCCAGGCGGCTGTGTGAGCTGGGGGAACGGCTGCGCTCAGAGAAAC	2524
QY	2565	AAGCCAGGCGGTATCAACAAGCTCCCTGTAGTTCGGGACTGTGATCAAAAGCAACCTGG	2624
Db	2525	AAGCCAGGCGGTATCAACAAGCTCTCTCTGTTCGGGACTGTGATCAAAAGCAACCTGTGG	2584
QY	2625	GTATAGCAGATGACAGACGCGCAACAAAAACCCACAGGGATGGCCGCAATGCATCACA	2684
Db	2585	GTATAGGCGCGGG---GCCACCCAATGTGTACCTCGGGGGCCACCCATGTCTCAACC	2641
QY	2685	CTGTGATACAGAGAGGAAACTGTGACGACATTATGTGTGGCTCTCCCCCCCAACACA	2744

Db 2642 CCAATGTGCACG - CCTGCAGGCTTGAGACTTGAGCCGCTGACTGCACCAAGCGCCC - CAGA 2699

Qy 2745 ACCCAGACTGTGAAGTGCATTCCTTAGAAGTCAAGT 2780

Db 2700 ACATACACTGTGAAGTCAATCTCCAGGCTTCAAAT 2735

```

RESULT 9
AAL53444
ID      AAL53444 standard; DNA; 3147 BP.

```

AC AAL53444

DT 12-DEC-2002 (first entry)

DE Type II transmembrane serine protease 1 coding DNA SEQ ID No 1.

KW Cytostatic; type-II membrane-type serine protease 7; MTSP7; malignancy; non-invasive; serine protease

KW malignant; enzyme; gene; ds.  
XY

OS Homo sapiens.  
XY

PN WO200272786-A2.  
XX

PD 19-SEP-2002.  
XX

PF 13-MAR-2002; 2002WO-US007903.  
XX

PR 13-MAR-2001; 2001US-0275592P.  
XX

PA (CORV-) CORVAS INT INC.  
XX

PI Madison EL, Ong EO;  
XX

DR WPI; 2002-732827,  
DR P-PSDB: 2A022929

XX  
PT New Transmembrane serine protease 7 (MTCB7) polypeptide for 3-dimensional

PT neoplastic diseases, monitoring tumor progress or therapeutic effectiveness or identifying MMR2 modulators for treating tumors

PT cancers.

PS Disclosure; Page 168-172; 184pp; English.  
XX

CC The invention relates to a purified single or two-chain polypeptide,  
CC which comprises the protease domain of a two-IT membrane-type serine  
CC

protease 7 (MTSP7) or its catalytically active portion. The polypeptid comprising MTSP7 is useful for detecting or diagnosing a neoplastic

CC disease, a pre-malignant lesion, a malignancy or other pathologic  
CC condition in a subject. This polymeride is also useful for monitoring

CC tumour (e.g. tumour of the breast, cervix, prostate, lung, ovary or  
CC colon) progress and/or therapeutic effectiveness. The inhibitor of the

polypeptide or the inhibitor of the polypeptide's zymogen form is useful for treating or preventing a neoplastic disease or tumour initiation

growth or progression, or a (pre-)malignant condition. The polypeptide or polynucleotide is also useful for identifying modulators of MTS27 which

may be used to treat cancers or tumours. This polynucleotide sequence represents the coding DNA for a protein of the type-II membrane-type

cc serine protease 1 relating to the invention  
xx

sequence 314 bp; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other;

Query Match	60.6%;	Score 1883.2;	DB 6;	Length 3147;
Best Local Similarity	81.2%;	Pred. No. 0;		

Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;

45 GATCGACCGCCAAACCATGGGTAGCAATCGGGCCGCAAGCCCGAGGGGCTCTCAG 104

D<sub>b</sub> 5 GAGCGCCTCGGGTACCATGGGAGCGATCGGGCCCGCAAGGCGGAGGGGGCCCGAAG 64

105 GACTTCGGCGGGA CTCAAGTACA ACTCCCGCTAGAGAACATGAATGGCTTTGAGGAG 164



Db 65 GACTTGCGCGGAGCTCAAGTACAACTCCCGCAGAGAAATGATGCTTGAGGAA 124  
Qy 165 GGTGTGAGATTCCTGCTCGGACGAACAATGCGAAGAAATGAGAGGAGGCCAGGCC 224  
Db 125 GGGGTGAGATTCCTGCGAGTCAACAGTCAAGAAAGTGGAAAGCATGGCCGGGGCCG 184  
Qy 225 TGGGTGTGTGTGTGGCAGTGTCTTCAAGTTCCTCTGTCTCTCTCTCTCTCTCTCT 284  
Db 185 TGGGTGTGTGTGTGGCAGTGTCTTCAAGTTCCTCTCTCTCTCTCTCTCTCTCTCT 244  
Qy 285 CTGATGTGCACTTCATTAATCGAAATGTGCGGGTTCAAAAAGTCTTCAATGAGCAT 344  
Db 245 CTGTGTGTGATTTGCAAGTACCGGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 304  
Qy 345 AGGATCAAAATGATGATCTTTCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 404  
Db 305 AGGATCAAAATGATGATTTTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGT 364  
Qy 405 AGCTGTGCGCAGGATGAGAGAGGCGGTGAGTGTGTGTGTGTGTGTGTGTGTGTGT 464  
Db 365 AGCTGTGCGCAGGATGAGAGAGGCGGTGAGTGTGTGTGTGTGTGTGTGTGTGTGT 424  
Qy 465 GGTCTTACCAAGAAGT 524  
Db 425 GGGCCCTACCAAGAAGT 484  
Qy 525 TACTGTGATGATTCAGATTCCTCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGT 584  
Db 485 TACTGTGATGATTCAGATTCCTCCCACTGTGTGTGTGTGTGTGTGTGTGTGTGT 544  
Qy 585 GTGAGAGGATTTGATTCATTTGCCACCCCGAGCAGGCGCATGAAATCTTGTGTGT 644  
Db 545 GAGAGAGGCTGATTCATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 604  
Qy 645 TCTGT 704  
Db 605 TCAGT 664  
Qy 705 AGTTTGT 764  
Db 665 AGCTTGT 724  
Qy 765 AACAGT 824  
Db 725 GACAGGCTGT 784  
Qy 825 GT 884  
Db 785 GT 844  
Qy 885 CTGT 944  
Db 845 CTGT 904  
Qy 945 GGCACCTTGT 1004  
Db 905 GGCACCTTGT 964  
Qy 1005 ACAGT 1064  
Db 965 ACAGT 1024  
Qy 1065 CCCAAGATGAGAGT 1124  
Db 1025 CCTGT 1084  
Qy 1125 TACTGT 1184  
Db 1085 TACTGT 1144  
Qy 1185 AACCGAAGT 1244

Db 1145 AACCGATGT 1204  
Qy 1245 GGTCTGT 1304  
Db 1205 GGTCTGT 1264  
Qy 1305 CAGTTGT 1364  
Db 1265 CAGTTGT 1324  
Qy 1365 TACAGGAGACCGGGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1424  
Db 1325 TACAGGAGACCGGGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1384  
Qy 1425 GGGATGT 1484  
Db 1385 GGGATGT 1444  
Qy 1485 GCAGATGCTGT 1544  
Db 1445 GCGCATGT 1504  
Qy 1545 AGT 1604  
Db 1505 AGT 1564  
Qy 1605 GGGAGCGAAGT 1664  
Db 1565 GGGAGCGAAGT 1624  
Qy 1665 GGGAGCGAAGT 1724  
Db 1625 GGGAGCGAAGT 1684  
Qy 1725 GAGAGGCTTGT 1784  
Db 1685 GAGAGGCTTGT 1744  
Qy 1785 CAAATGT 1844  
Db 1745 CTGATGT 1804  
Qy 1845 GATGT 1904  
Db 1805 GAGGCTTGT 1864  
Qy 1905 GT 1964  
Db 1865 GT 1924  
Qy 1965 CTGGGCGAGGGCCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2024  
Db 1925 CTGGGCGAGGGCCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1984  
Qy 2025 GCTATGT 2084  
Db 1985 GCACTGT 2044  
Qy 2085 TTCTGT 2144  
Db 2045 TTCTGT 2104  
Qy 2145 AAAGT 2204  
Db 2105 AAAGT 2164  
Qy 2205 CTGT 2264  
Db 2165 CTGT 2224  
Qy 2265 GCTTGT 2324  
Db 2225 GCTTGT 2284

QY	2325	GAGGGAGAGTACCGGAGCGCGTATCTCTGCAGAAAGGCTGAGATTCGTGTATCAACCAAGC	2384
Db	2285	TATGGAGGCATCGGCGCGGTGATCTTCGACAAAAGGCTGAGATCCGCGTATCAACCAAGAC	2344
QY	2385	ACCTGTGAGGACCTCATGCGCAGCAGATCACCCCAAGATGATGTGTGGGCTTCTC	2444
Db	2345	ACCTCGAGAACTCTCGTCGCCAGAGATACGCCCGCATGATGTGCTGGCTTCTC	2404
QY	2445	AGTGGGGGTGTGGACTCTCTGCCAGGGTACTCTGTGGCCCCCTTGTCAAGCGCGAGAA	2504
Db	2405	AGCGCGCGCGTGGACTCTCTGCCAGGGTATTTCCGGGGGACCCCTGTCCAGCGTGAAGCG	2464
QY	2505	GATGGCGCAATGTTCCAGGCTGTGTGGTGGAGCTGGGGTGAAGGTGTGGCTCAAGAGAAC	2564
Db	2465	GATGGCGCGATCTTCCAGGCGGGTGTGTGGTGGAGCTGGGGAGACGGTGTGGCTCAAGAGAAC	2524
QY	2565	AAGCCAGGCGTGTACCAAAAGGCTCCCTTGTAGTTGGGATCGAGATCAAAAGACACACTGGG	2624
Db	2525	AAGCCAGGCGGTGTACCAAAAGGCTCCCTGTGTTTGGGACTGGATCAAAAGAACACTGGG	2584
QY	2625	GTATAGCAGCATGGACGACAGCGCGACCAACCAACCCACAGGGATGCCCCAATGCACA	2684
Db	2585	GTATAGGGGCGCGG---GCAACCAATGTGTACACTGCGGGGCCAACCCATCGTCCACC	2641
QY	2685	CCTGGATACAGAGAGAGAACTGACGACATTTATGTGTGGCTCTCCCCCCCCCAACACA	2744
Db	2642	CCAATGTGTACG-CCTGCAAGGCTGGAGACTGGAACCGGTGACTGACCAACGAGCCC-CCAGA	2699
QY	2745	ACCCAGACTGTGAATCTGCATCTTTAGACTAGAGT	2780
Db	2700	ACATACACTGTGAATCAATCTCCAGGGCTTCCAAT	2735

	RESULT 10
XX	AAL53445
ID	AAL53445 standard; DNA; 3147 BP.
XX	
AC	AAL53445;
XX	
DT	12-DEC-2002 (first entry)
XX	
DB	Type II transmembrane serine protease 1 domain DNA SEQ ID No 3.
XX	
KW	Cysteatic; type-II membrane-type serine protease 7; MTSPT7; malignancy;
KM	neoplastic disease; pre-malignant lesion; tumour; zymogen form; cancer;
Kw	malignant; enzyme; gene; ds.
XX	
OS	Homo sapiens.
XX	
PN	MO200272786-A2.
XX	
PD	19-SEP-2002.
XX	
PF	13-MAR-2002; 2002WO-US007903.
XX	
PR	13-MAR-2001; 2001US-0275592P.
XX	
PA	(CORV-) CORVAS INT INC.
XX	
P1	Madison EL, Ong EO;
XX	
DR	WPI; 2002-732827/79.
DR	P-PADB; AAO22930.
XX	
XX	
PT	New transmembrane serine protease 7 (MTSP7) polypeptide for diagnosing
PT	neoplastic diseases, monitoring tumor progress or therapeutic
PT	effectiveness, or identifying MTSP7 modulators for treating tumors or
PS	cancers.
PS	Disclosure, Page 174-176, 184pp; English.
CC	The invention relates to a purified single or two-chain polypeptide.

CC Which comprises the protease domain of a type-II membrane-type serine  
CC protease 7 (MTSP7) or its catalytically active portion. The polypeptide  
CC comprising MTSP7 is useful for detecting or diagnosing a neoplastic  
CC disease, a pre-malignant lesion, a malignancy or other pathologic  
CC condition in a subject. This polypeptide is also useful for monitoring  
CC tumour (e.g. tumour of the breast, cervix, prostate, lung, ovary or  
CC colon) progress and/or therapeutic effectiveness. The inhibitor of the  
CC polypeptide or the inhibitor of the polypeptide's zymogen form is useful  
CC for treating or preventing a neoplastic disease, or tumour initiation,  
CC growth or progression, or a (pre-)malignant condition. The polypeptide o  
CC polynucleotide is also useful for identifying modulators of MTSP7, which  
CC may be used to treat cancers or tumours. This polynucleotide sequence  
CC represents the coding DNA for a protein domain of the type-II membrane-  
CC type serine protease 1 relating to the invention

Query Match	Best Local Similarity	Score	DB	Length
60.6%	81.2%	1883.2	6	3147
Pred. Local Similarity	Pred. No. of			
Matches 2233	Conservative	0	Mismatches 508	Indels 5
			Gaps	3
45	GATCGGACCGCCAAAACCATGGGTGCAATGCGGGGCGCCAAAGCGCGAGGGGGCTCTCAG	104		
5	GAGCGGCGCTCGGGGATACATGGGAGGCGATTCGGGCGCCAAAGGGGAGGGGCCCGAAG	64		
105	GACTTCGGCGCGGGACTTAAGTACAACCTCCGGCTTAGAAATGATGCTTTGAGGAG	164		
65	GACTTCGGCGCGGGACTTAAGTACAACCTCCGGCAAGGAAAGTAAATGCTTTGAGGAA	124		
165	GATGTCGAGTTCCTTCCTGCGCAACAATGCAAGAAAGTGAAGCGAGCGCCCGAGGGC	224		
125	GCGCTGGAGTTCTCTCCAGTACAACCTCAAGAAAGTGAAGAAAGCATGCGCCGGGGGCG	184		
225	TGGGTGTCGTGTGGCACTGCTGTCAAGTTCCCTTCCTCTCTCCCTCATGAGTGGCTTG	284		
185	TGGGTGTCGTGTGGCACTGCTGTCAAGTTCCCTCTCTCTTCCTTCCTTCGTGGGATGGCTTC	244		
285	CTGGTGTGGCACTTCATTAATCGAAATGTGGGCTTCAAAAAGTCTTCAATGCGCAATG	344		
245	CTGGTGTGGCACTTCATTAATCGAAATGTGGGCTTCAAAAAGTCTTCAATGCGCAATG	304		
345	AGGATCAACAATGAGATCTTCTGATGTCGATAGAAATCCTCACTCCAGACGATTTATC	404		
305	AGGATCAACAATGAGATTTTGTGATGTCCTACAGAACTTCAACTCCTCACTGAATTTGTA	364		
405	AGCTGCGCAAGTGTGAAGAGGCGCTGAAGCTGCTGTAACAATGAATGCTGTCTGT	464		
365	AGCTGCGCAAGTGTGAAGAGGCGCTGAAGCTGCTGTAACAAGGATGCCATTTCTGT	424		
465	GGTCCCTTACCAAGAAAGTGGCTGTAACTGCTTCACTGAGGGCATGTCATGCTTAC	524		
425	GGCCCTTACCAAGAAAGTGGCTGTAACTGCTTCACTGAGGGCATGTCATGCTTAC	484		
525	TACTGTGACAGTTCAGCATCCTCCCAACCTGCGCAAAAGGTTGATTCGCGCATGCT	584		
485	TACTGTGACAGTTCAGCATCCTCCCAACCTGCGCAACCTGCGCAAGGCTGTCATGCT	544		
585	GTGAGAGGAGTGTGAACATTTGCCAAGCCCGAGCAAGCGGCACTGAATCTTGTGCTACA	644		
545	GAGAGAGGAGTGTGAACATTTGCCAAGCCCGAGCGGCTGCTTGAAGTCTTTGTGTGACC	604		
645	TCTGTGAGGCTTCCCATTTGACCCCAAGATGTGAGAGGACTACAGCAACAGCTGC	704		
605	TCAATGTGCTTTCCCATCGAATCTCAAAACATTAAGAGAGCCAGGCAACAGCTGC	664		
705	AGTTTGCCTGATGCGCATGCTGTGAGAGTGAACGCTTCACTACCTCTGCTTCCC	764		
665	AGTTTGCCTGATGCGCGCGGAGTGTGAGCTGATGCTTCAACAGCGCGCTTCCCT	724		
765	AACGATCCCTAACCGGCGCATGCGCTGTGCAATGCTGCTGTGGGGGAGCGCACTCT	824		
725	GACGCGCCCTAACCGGCTGATGCTGTGCAATGCTGCTGTGGGGGAGCGCACTCA	784		

QY	825	GTGCTAGACCTCA	CTTTCCGAACCTTTG	ATGTGCTCCCTG	TATGAGATGGCACTGAC	884																
Db	785	GTGTGTAGGCTTCA	CTTCCGACCTTTGAC	CTTGCTCTGCGACGACGCGGCA	CGAC	844																
QY	885	CTGATCA	CCCTGTATGATAG	AGCTGAGCCCATGAA	CCCA	CGCTGTGTG	944															
Db	845	CTGGTGA	CGGTATCA	CA	CCCTGAGCCCATGAG	CCCA	CGCTGTGTG	904														
QY	945	GGCACCTTCTCA	CCCTCTCA	CACTTTCTCTCTCC	CGAAGCGTCTCTTGTG	1004																
Db	905	GGCACCTTAC	CCCTCTCTCA	CACTTTCTCTCTCC	CGAAGCGTCTCTTGTG	964																
QY	1005	ACCGTGA	TAACCA	TA	CTGACCCGCGACAT	CTGTGAGGCACTTCTCTCA	CGTGS	1064														
Db	965	ACACTGA	TAA	CA	CA	CTGAGCGCGGCA	CTCCGCTTTAGGCGACCTTCTTCA	CGTGS	1024													
QY	1065	CCCAAGATGA	GCAC	AGTGTGGCGCTTTT	TGATGAC	CCCAAGGACATTTAG	CA	CGCC	1124													
Db	1025	CTTGA	GTATGAC	AGCTGTGGAGGCGCTT	ATGTAA	CCCAAGGACATTTCA	CA	CGCC	1084													
QY	1125	TACTATCA	AGG	CCACTAC	CCG	CCCAATCA	CTGCA	CA	TGAAATATCA	AGTGGCCCA	1184											
Db	1085	TACTATCA	AGG	CCACTAC	CCG	CCCAATCA	CTGCA	CA	TGAAATATCA	AGTGGCCCA	1144											
QY	1185	AA	CCGGAACG	GAAGGAGGCGCTTCA	AACTCTTATCTGTGTG	ACCCCA	CGTAC	CA	AGTGS	1244												
Db	1145	AA	CCGGAACG	GAAGGAGGCGCTTCA	AACTCTTATCTGTGTG	ACCCCA	CGTAC	CA	AGTGS	1204												
QY	1245	GGCTCTG	CA	CA	GA	ACTATGTG	AGATCA	CA	CGGGA	AGA	ATCTGCGGTG	AGGTCC	1304									
Db	1205	GGCA	CTG	CCCA	AG	ACTATGTG	AGATCA	CA	CGGGA	AGA	ATCTGCGGTG	AGGTCC	1264									
QY	1305	CAGTTTGTG	TGAG	AGAC	CA	CA	GA	AGATTA	CA	GTCA	CTTCCATTTCT	GTATCA	CTCG	1364								
Db	1265	CAGTTTGTG	TGAG	AGAC	CA	CA	GA	AGATTA	CA	GTCA	CTTCCATTTCT	GTATCA	CTCG	1324								
QY	1365	TAC	AC	CGA	CA	CGG	GGTCTT	AC	GTG	AG	AACTCTCT	AC	GA	CTCA	CGA	CCGCGT	CGCA	1424				
Db	1325	TAC	AC	CGA	CA	CGG	GGTCTT	AC	GTG	AG	AACTCTCT	AC	GA	CTCA	CGA	CCGCGT	CGCA	1384				
QY	1425	GGGATG	TTTCA	TGTG	CA	GA	ACTG	CA	CGTGC	ATCC	GAA	AGAACTG	CGCTG	CA	CGTGS	1484						
Db	1385	GGGATG	TTTCA	TGTG	CA	GA	ACTG	CA	CGTGC	ATCC	GAA	AGAACTG	CGCTG	CA	CGTGS	1444						
QY	1485	GC	AG	CTG	CCCG	GA	TTATAT	GTATG	ATG	AGCTTACTG	CCGATG	CA	TGCC	CA	CCCA	CA	AGTTC	1544				
Db	1445	GC	AG	CTG	CCCG	GA	TTATAT	GTATG	ATG	AGCTTACTG	CCGATG	CA	TGCC	CA	CCCA	CA	AGTTC	1504				
QY	1545	AC	GTG	CA	AAAA	CA	AGTTCTG	CA	AG	CGCCCTCTT	GTG	GTG	CA	GTG	CA	AGTCA	AG	1604				
Db	1505	AC	GTG	CA	AAAA	CA	AGTTCTG	CA	AG	CGCCCTCTT	GTG	GTG	CA	GTG	CA	AGTCA	AG	1564				
QY	1605	GGGACG	GA	AGTGA	CG	AG	AGGAGG	AGCTG	CTCTG	TG	GG	AGTTCA	CA	AGTGTTC	CA	AT	1664					
Db	1565	GGGACG	GA	AGTGA	CG	AG	AGGAGG	AGCTG	CTCTG	TG	GG	AGTTCA	CA	AGTGTTC	CA	AT	1624					
QY	1665	GGGAA	GTGTCTTCC	CA	AG	CCGA	AGTGTAT	GGGAG	AGCA	CTGTG	GA	TGGGTCT	1724									
Db	1625	GGGAA	GTGTCTTCC	CA	AG	CCGA	AGTGTAT	GGGAG	AGCA	CTGTG	GA	TGGGTCT	1684									
QY	1725	GAC	GAG	GCTTCA	TGTG	CA	CG	GTGA	ATG	CGCTTCTT	CA	CA	AA	TAT	AT	CA	CG	GTGC	1784			
Db	1685	GAC	GAG	GCTTCA	TGTG	CA	CG	GTGA	ATG	CGCTTCTT	CA	CA	AA	TAT	AT	CA	CG	GTGC	1744			
QY	1785	CA	AA	AT	GTG	CCCTGT	GTCTG	AG	CA	AGGCA	CA	CCCTG	ATG	TG	GA	GA	AG	CA	CG	ATG	AGC	1844
Db	1745	CT	CA	AT	GTG	CCCTGT	GTCTG	AG	CA	AGGCA	CA	CCCTG	ATG	TG	GA	GA	AG	CA	CG	ATG	AGC	1804
QY	1845	GATG	CT	CG	AT	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	1904
Db	1805	GATG	CT	CG	AT	G	A	A	A	A	A											

Db		1865	GTTGTTTTGGGGCACCGAATGTGGAGTGAAGGGGAGAGTGggcccttgacagatgaagcgttcattgct	1924
OY		1965	CTGGGCCACAGGGCCACTTGTGTGGGGCCTGCCTCATCTCTCTGACTGGCTGTGTTCTTCGCA	2024
Db		1925	CTGGGCCACAGGGCCACAATCTGCGGTGCTTCCCTCATCTCTCCCAACTGGCTGTGCTTCGCG	1984
OY		2025	GCTCATTTGCTTTTAGAGATGACAAAATTTCAGAAGTCTCAGACTCACGATGTGGACGGCC	2084
Db		1985	GCAACAAGCTTAACATCGATGACAGAGAGTTACAGGTAAGCTACAGCCCCACGACGTGGACGGCC	2044
OY		2085	TTCCTGGGCTCTGCTGGACACAGACAAAGCGAGTGCCCTCTGGGGGTGACAGAGCTGAAGCTC	2144
Db		2045	TTCCTGGGCTTTGACACAGACCAGAGCGACGCGCCCTGGGGGTGACAGAGCGACGGCTC	2104
OY		2145	AAAAGTATCATCAACCCACCCCTTCTCTCATGATTTACCTTGACTATGACATGGCCTTG	2204
Db		2105	AAGGAGCATCAATCTCCCAACCCTTCTTCAATGACTTCACCTTGACATAATGACATGGCGCTG	2164
OY		2205	CTGAGAGCTGGAGAAAGTGGGTGGAGTAAGACAACGGTCGGCGCCCAATCTGCTGGCTGAT	2264
Db		2165	CTGGAGCTGGAGAAACCGGAGAGATTAAGCTTCAATGGTCGGGCCCAATGTCTTGCCGGAC	2224
OY		2265	GCTAACCATGATCTTCCCTGCTGGAGAAAGGCATCTGGGTCACAGGCTGGGGGCACACAAAA	2324
Db		2225	GCTCCTCATGTTCTTCTCTGCGGCAAGGCCATCTGGGTCAAGGGGTGGGGACACACCCAG	2284
OY		2325	GAGGAGAGTTACCGAGGCGCTGATCTCTGACAAAGGGTGAATCCGTTCATCAACAGAC	2384
Db		2285	TATGAGAGCACTGGCGGGCTGATCTCTCAAAAAGGGTGAATCCGCTCATCAACAGACC	2344
OY		2385	ACCTGTAGAGACCTCATATGCCGACACATCAACCCCACAGAAAGATGTGTGTGGGTTCTCTC	2444
Db		2345	ACCTGCAGAAACCTCTCTGCGGACAGATCACGCCGCGATGATGTGCTGTGGCTTCTTC	2404
OY		2445	AGTGGGGGTGTGGAAGCTCTGCGCAGGGTGAATCTGGTGGCCCCCTTGTCAAGCGCGAGAAA	2504
Db		2405	AGCGGCGGCGTGGACTCTGCGACAGGTGATTTCCGGGGGACCCCTGTCCAGGGTGAAGGCG	2464
OY		2505	GATGGGCGAATGTTCCAGGCTGTGTGTGATGCTGGGSTGAAGCTGCGCTCAAGAGAAC	2564
Db		2465	GATGGGCGAATCTTCCAGGCGGGTGTGTGAGCTGGGGACAGCGCTGCGCTCAAGAGAAC	2524
OY		2565	AAAGCAGGCGCTGTACACAAAGGCTCCCTGTAGTTGGGGACTGGATCAAAAGACACACTGGG	2624
Db		2525	AAAGCAGGCGCGTGTACACAAAGGCTCCCTGTGTTGGGGACTGGATCAAAAGAGAACACTGGG	2584
OY		2625	GTAATAGCAGATGAGACAGACAGCGGACACAAAACCCACAGAGGATGGCCGACATGACA	2684
Db		2585	GTAATAGGGGGCCGGG---GCCAACCCAATGTGTAACCTGCGGGGGCCACCCATGTGTCAAC	2644
OY		2685	CCTGGATTCAGAGAGAGAAACA CTGACGACATTTATGTGTGGCTTCCGCCCTCCCAACACA	2744
Db		2642	COAGTGTCAAG-CCTGCAGGCTGGAGACTGGACGCTGACTGACGACGAGCGCC-CCAGA	2694
OY		2745	AACCAGACTGTGAAGTGCATCTTGTGAGACTCAGAGT	2780
Db		2700	ACATACACTGTGAATCAATCTCCAGGGGCTCCAAAT	2735
<hr/>				
RESULT 11				
ABZ22450				
ID	ABZ22450	standard; cDNA, 3147 BP.		
XX	ABZ22450;			
XX	AC	(first entry)		
XX	DT	24-MAR-2003		
XX	XX	Human membrane-type serine protease MTSPI encoding cDNA SEQ ID NO:1.		
XX	XX	Human; membrane-type serine protease; enzyme; MTSPI0; cytosolic;		
KW	type-II membrane-type serine protease; neoplastic disease; tumour; MTSPI;			
KW	multipase; gene; ss.			

```

XX OS Homo sapiens.
XX Key Location/Qualifiers
XX Key 23..2590
XX FT /tag= a
XX FT /product= "membrane-type serine protease MTSPI"
XX FT /note= "also known as matrilysin"
XX PN WO200292841-A2.
XX PD 21-NOV-2002.
XX PR 14-MAY-2002; 2002WO-US015332.
XX PR 14-MAY-2001; 2001US-0291001P.
XX PA (CORV-) CORVAS INT INC.
XX PI Madison EL, Yeh J;
XX DR MPI; 2003-129309/12.
XX DR P-PSDB; ABB56619.
XX PT New polypeptides comprising the protease domain of a type-II membrane-
XX PT type serine protease (MTCPI0), or its mutants, useful for diagnosing
XX PT neoplasms or malignancies, or for screening for MTCPI0 inhibitors for
XX PT treating such diseases.
XX PS
XX PS Example 1; Page 177-181; 1989p; English.
XX CC The present invention describes a polypeptide comprising a purified
XX CC single or two chain polypeptide, which comprises the protease domain of a
XX CC type-II membrane-type serine protease (MTSP10) or its catalytically
XX CC active portion, or a mutant of it, where up to 50 % of the amino acids
XX CC are replaced with another amino acid, and the resulting polypeptide is a
XX CC single chain or two chain polypeptide that has a catalytic activity of at
XX CC least 1-10 % of the unmutated polypeptide. MTSPI0 has cytostatic
XX CC activity. The polypeptide containing the protease domain of the MTSPI0 is
XX CC useful for detecting a neoplastic disease, and for diagnosing the
XX CC presence of a pre-malignant lesion, a malignancy, or other pathologic
XX CC condition in a subject, or monitoring tumour (e.g. breast, cervix,
XX CC prostate, lung, ovary or colon tumour) progression and/or therapeutic
XX CC effectiveness. An inhibitor of the polypeptide containing the protease
XX CC domain of MTSPI0 is useful for treating or preventing neoplastic disease
XX CC in a mammal. An inhibitor of the activation or cleavage of the zymogen form
XX CC of the MTSPI0 polypeptide is useful for inhibiting tumour initiation,
XX CC growth or progression, or treating (pre-)malignant conditions of the e.g.
XX CC breast, cervix, prostate, lung, ovary or colon. The present sequence
XX CC encodes human MTSPI (also known as matrilysin), which is used in an
XX CC example from the present invention
XX SQ Sequence 3147 BP; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other;
Query Match 60.6%; Score 1883.2; DB 7; Length 3147;
Best Local Similarity 81.2%; Pred. No. 0;
Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;
QY 45 GATGCGACCGCCAAACCATGGGTAGCAATGGGGCCGCAAGCGGGGGGCTCTAG 104
Db 5 GAGCGGCTCGGGGATCCATGGGAGCGATCGGCGCCGCAAGGGGGGGCCCGAAG 64
QY 105 GACTTCGGCGCGGAGCTCAAGTCAACTCCGGCTAGAGAAATGATGCTTTGAGAG 164
Db 65 GACTTCGGCGCGGAGCTCAAGTCAACTCCGGCGAGAAAGTAAATGCTTGAAGAA 124
QY 165 GGTGTGAGTTCCTGCTCCGCAACAATGCGCAAGAAAGTGAAGACCGAGCCCGAGGCGC 224
Db 125 GCGGTGAGTTCCTGCTCCGCAACAATGCGCAAGAAAGTGAAGAAACATGAGCGCCGAGGCGC 184
QY 225 TGGGTGAGTTCCTGCTCCGCAAGTGTGCTTCAAGCTTCTTGTGCTTCTGAGGCTG 284
Db 185 TGGGTGAGTTCCTGCTCCGCAAGTGTGCTTCTTGTGCTTCTGAGGCTG 244

```

```

QY 285 CTGGTGTGCACTTCATTCAGGAATGGCGGTTCAAAAAGTCTTCAATGGCCATCTG 344
Db 245 CTGGTGTGCACTTCATTCAGGAATGGCGGTTCAAAAAGTCTTCAATGGCCATCTG 304
QY 345 AGGATCACAAAATGAGATCTTCTTGTGATGGATGAGAACTCCACCTCCACAGATTATC 404
Db 305 AGGATCACAAAATGAGATCTTCTTGTGATGGATGAGAACTCCACCTCCACAGATTATC 364
QY 405 AGCTTGGCCAGCGAGTGAAGAGGCGTGAAGCTGCTGTACATGAATCCCTGTCTG 464
Db 365 AGCTTGGCCAGCGAGTGAAGAGGCGTGAAGCTGCTGTACATGAATCCCTGTCTG 424
QY 465 GGTCCCTACCAAGAAGTGGGTGTAACCTGCTTGAAGAGGAGGAGTCAATCGGCTAC 524
Db 425 GGTCCCTACCAAGAAGTGGGTGTAACCTGCTTGAAGAGGAGGAGTCAATCGGCTAC 484
QY 525 TACTGTGCAAGTTCAGATCCCGCCACACTGGCAGAAAGGTTGATGCGCCATGAGT 584
Db 485 TACTGTGCAAGTTCAGATCCCGCCACACTGGTGAAGAGGCGCCATGAGT 544
QY 585 GTGAGCGAGTTGTAACTTGCACCCGAGACCGGCACTGAATCCTTGTGTTAA 644
Db 545 GAGGAGCGGTGATGATGCTGCTGCGCCGCGGCGGCTCTGTAAGTCTTGTGTACAC 604
QY 645 TCTGTGTGAGCTTCCCGCCATGACCCGAGAAATGTGACAGAGACTCAGAGCAACAGCTG 704
Db 605 TCAGTGTGTGCTTTCCTCCACGCACTTCGCAAAACATGACAGAGACCCAGACACAGCTG 664
QY 705 AGTTTTCCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 764
Db 665 AGTTTTCCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 724
QY 765 AACAGTTCCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 824
Db 725 GACAGCTTCCTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 784
QY 825 GTGCTGAGCTTACCTCCGAGCTTGTGATGATGATGATGATGATGATGATGATGATG 884
Db 785 GTGCTGAGCTTACCTCCGAGCTTGTGATGATGATGATGATGATGATGATGATGATG 844
QY 885 CTGCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 944
Db 845 CTGCTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 904
QY 945 GGCACCTTTCACCTTCTTCAACCTGATGATGATGATGATGATGATGATGATGATGATG 1004
Db 905 GGCACCTTTCACCTTCTTCAACCTGATGATGATGATGATGATGATGATGATGATGATG 964
QY 1005 ACGCTGATTAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1064
Db 965 ACACGTGATTAACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1024
QY 1065 CCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1124
Db 1025 CCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1084
QY 1125 TACTATCCAGGCTTACCTCCGAGCTTCAACCTGATGATGATGATGATGATGATGATGATG 1184
Db 1085 TACTATCCAGGCTTACCTCCGAGCTTCAACCTGATGATGATGATGATGATGATGATGATG 1144
QY 1185 AACCGAAGTGAAGAGTGGCTTCAACCTTCTTGTGATGATGATGATGATGATGATGATG 1244
Db 1145 AACCGAAGTGAAGAGTGGCTTCAACCTTCTTGTGATGATGATGATGATGATGATGATG 1204
QY 1245 GGTCTTCGACCAAGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1304
Db 1205 GGTCTTCGACCAAGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1264
QY 1305 CAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1364
Db 1265 CAGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1324

```

```

QY 1365 TACAGGAGACCGGGTTCCTAGTAGTACTCTCTACAGACTCCAGACGCCGTGCCA 1424
DB 1325 TACACGACACCGGGCTTCTTAGCTGAATCTCTCTACAGACTCCAGACGCCGTGCCA 1384
QY 1425 GGGATTTGATGTGCAAGACTGACCGGTGATCCGAAGAACTGGCTGCCAGCGCTGG 1484
DB 1385 GGGCAATTCACGTGCGGACGCGGGGGGTGTATCCGGAAGGAGCTGGCGCTGTGATGGCTGG 1444
QY 1485 GCAGACTGCCCCGATTAATAGTAGTACGCTTACTGCCGATGCAATGCCACCCACGATTC 1544
DB 1445 GCGGACTGCAACCGACACAGCGATGAGCTTCAGCTGACGTCACGCCCGGCCACGAGTTC 1504
QY 1545 AGGTGCAAAAACCAAGTTCTGCAAGCCCTCTCTGGGTCTGTGACAGTGTCAACGACTGT 1604
DB 1505 AGGTGCAAGAACAGTTCTGCAAGCCCTCTCTGGGTCTGTGACAGTGTGAACGACTGTG 1564
QY 1605 GGGGACGAGATGACAGAGAGGAGGCTGACGCTGTCTGTGGAGATTCAAGTGTCCAAAT 1664
DB 1565 GGAAGACACAGCGACGAGCGGGGTGCAATTGTCCGGCCACAGCTTCAGGTGTCCAAAT 1624
QY 1665 GGGAAATGTCTCCCTCAGAGCCGAAAGTGTATGGGAAGACAACTGTGAGATGGGTCT 1724
DB 1625 GGGAAATGTCTCTCGAAGAACCCAGAGTGTGAATGGGAAGACAACTGTGGGACGAGTTC 1684
QY 1725 GACGAGGCTTCATGTGACAGCGTGAATGTCTCTGTGACCAAAATATACCTACCGCTGC 1784
DB 1685 GACGAGGCTCTCTGCCCCAGAGTGAAGTGTCTGTGACCAAAACAACTTCACCGCTGC 1744
QY 1785 CAAATGTGCTCTGTCTGACCAAGGCAACCCGTAGTGTATGGGAAGACGAGCTGTAGC 1844
DB 1745 CTCATGTGGCTCTGCTTGACCAAGGCAACCCGTAGTGTGACGGAAGAGAGACTGTAGC 1804
QY 1845 GATGTCTCCGATGAGAAAACTGTGACTGTGGCTGTGATCTTTTACCAACAGCGCTGC 1904
DB 1805 GAGCGCTCAGATGAGAACTGTGACTGTGGCTGTGGCTATTCAGAGAACAGCGCTGT 1864
QY 1905 GTGGTTGTGGACAGATGCGGACGAGAGGAGGCGCTGTGGAGTGAAGCTGTCAAGCC 1964
DB 1865 GTTGTGTGGGACAGATGCGGATGAGGAGGAGTGGCCCTGTGGAGTGAAGCTGTGATGCT 1924
QY 1965 CTGGGCCAGGAGCCATTGTGTGGGGCTCGTCAATCTCTCTGACCAAAATATACCTACCGCTGC 2024
DB 1925 CTGGGCCAGGAGCCATTGTGTGGGGCTCTCCCTCAATCTCTCCCACTGGCTGTGTGCGC 1984
QY 2025 GCTATTTGCTTTACGATGAGCAAAATTTTCAAGTATCAGACTACAGATGTGAGCGCC 2084
DB 1985 GCACACTGTCAATCAATGACAGAGATTCAGTACTGAGCCCAAGCAATGTGAGCGCC 2044
QY 2085 TTCTGTGGGTCTGTGACAGAGAGAGGAGGAGTGTCTGTGGGTGACAGAGCTGMACTC 2144
DB 2045 TTCTGTGGGTCTGTGACAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2104
QY 2145 AAACGTATCATCACTACCCACTTCTTCAATGATTTCACTTTGACATGATGAGCTGTG 2204
DB 2105 AAGGCAATCATCTCCACCCCTTCTTCAATGATTTCACTTTGACATGATGAGCTGTG 2164
QY 2205 CTGAGAGCTGAGAGAGTGTGTGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2264
DB 2165 CTGAGAGCTGAGAGAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2224
QY 2265 GCTACCCATGTCTTCCCTGTGGCAAGGCACTGTGGGTCAAGGCTGTGGGGGACACAAAA 2324
DB 2225 GCTTCCATGTCTTCCCTGTGGCAAGGCACTGTGGGTCAAGGAGGAGGAGGAGGAGGAGG 2284
QY 2325 GAGGAGGTACCGAGCGCTGATCTGTGAGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2384
DB 2285 TATGAGAGCACTGGGCGGCTGATCTGTCAAAAAGGAGGAGGAGGAGGAGGAGGAGGAGG 2344
QY 2385 ACCTGTGAGAGGCTTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2444
DB 2345 ACCTGTGAGAGGCTTCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2404
QY 2445 AGTGGGGGTGTGAGACTCTGTGCGAGGAGTACTGTGTGGCCCTTGTCAAGGCGGAGAAA 2504

```

```

DB 2405 AGGCGGCGGTGAGACTCTGCGAGGAGGTGATCCGAGGAGACCCCTGTCCAGCGTGGAGGCG 2464
QY 2505 GATGGGCGAATGTTCCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2564
DB 2465 GATGGGCGAATGTTCCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2524
QY 2565 AAGCCAGGCGGTGTACACAAGGCTCCCTGTAGTTCGAGACTGTATCAAGAGCACTGTGG 2624
DB 2525 AAGCCAGGCGGTGTACACAAGGCTCCCTGTAGTTCGAGACTGTATCAAGAGCACTGTGG 2584
QY 2625 GTATGACAGATGACAGACAGCGGACCAACAAACCCACAGGAGATGCCGATGACCA 2684
DB 2585 GTATGAGGCGCGGG--GCCACCCAAATGTGTACACTGTGCGGGGCCAATGTGTCCACC 2641
QY 2685 CCGTATGACAGAGAGAGACACATGACATTTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2744
DB 2642 CCGTGTGACAG--CTGTGAGGCTGTGAGACTGTGACCGTGTACTGTGACAGGCGCC--CCAG 2699
QY 2745 ACCGAGCTGTGAACTGATCTTGTAGACTCAGAGT 2780
DB 2700 ACATCACTGTGAACTGATCTTGTAGACTCAGAGT 2735

```

RESULT 12  
ABZ22451  
ID ABZ22451 standard; cDNA; 3147 BP.

XX ABZ22451;

DT 24-MAR-2003 (first entry)

XX Human MTSPI protease domain encoding cDNA SEQ ID NO:3.

XX Human; membrane-type serine protease; enzyme; MTSPI0; cytostatic;  
KW type-II membrane-type serine protease; neoplastic disease; tumour; MTSPI;  
KW matrix; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers  
FT CDS 1865..2530  
FT /tag= a  
FT /partial  
FT /product= "MTSPI protease domain"  
FT /note= "no start codon given"

XX W0200292841-A2.

XX 21-NOV-2002.

XX 14-MAY-2002; 2002WO-US015332.

XX 14-MAY-2001; 2001US-0291001P.

XX (CORV-) CORVAS INT INC.

XX Madison EL, Yeh J;

XX WPI; 2003-129309/12.

XX P-PSDB; ABP56620.

XX New polypeptides comprising the protease domain of a type-II membrane-  
PT type serine protease (MTCPI0), or its mutants, useful for diagnosing  
PT neoplasms or malignancies, or for screening for MTCPI0 inhibitors for  
PT treating such diseases.

XX Example 1; Page 183-185; 1989p; English.

XX The present invention describes a polypeptide comprising a purified  
CC single or two chain polypeptide, which comprises the protease domain of a  
CC type-II membrane-type serine protease (MTSPI0) or its catalytically  
CC active portion, or a mutant of it, where up to 50 % of the amino acids

CC are replaced with another amino acid, and the resulting polypeptide is a  
 CC single chain or two chain polypeptide that has a catalytic activity of at  
 CC least 1-10 % of the unmutated polypeptide. MTS10 has cytosolic  
 CC activity. The polypeptide containing the protease domain of the MTS10 is  
 CC useful for detecting a neoplastic disease, and for diagnosing the  
 CC presence of a pre-malignant lesion, a malignancy, or other pathologic  
 CC condition in a subject or monitoring tumour (e.g. breast, cervix,  
 CC prostate, lung, ovary or colon tumour) progression and/or therapeutic  
 CC effectiveness. An inhibitor of the polypeptide containing the protease  
 CC domain of MTS10 is useful for treating or preventing neoplastic disease  
 CC in a mammal. An inhibitor of the activation cleavage of the zymogen form  
 CC of the MTS10 polypeptide is useful for inhibiting tumour initiation,  
 CC growth or progression, or treating (pre-)malignant conditions of the e.g.  
 CC breast, cervix, prostate, lung, ovary or colon. The present sequence  
 CC encodes the protease domain of human MTS1 (also known as matrilase),  
 CC which is used in an example from the present invention

XX Sequence 3147 BP; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other;

Query Match 60.6%; Score 1883.2; DB 7; Length 3147;  
 Best Local Similarity 81.2%; Pred. No. 0;  
 Matches 223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;

45 GATCGAGCGCCAAACCATGGTAGCAATCGGGGCGGCAAGGCGGAGGGGCTCTCAG 104  
 5 GAGCGGCTTCGGGGTACCATGGGAGGATCGGGCGGCAAGGCGGAGGGGCGGAGG 64  
 105 GACTTGGGCGGGGACTCAAGTACAACTCCGCGTAGAGAACTGAATGCTTTGAGAG 164  
 65 GACTTGGGCGGGGACTCAAGTACAACTCCGCGTAGAGAAATGATGGCTTGGAGAA 124  
 165 GGTGTGGAATTTCTGCTCGCAACAATGCCAAGAAAGTGAAGAGGAGGCGGCGGCG 224  
 125 GGCCTGGAATTTCTGCTCGCAACAATGCCAAGAAAGTGAAGAGGAGGCGGCGGCG 184  
 225 TGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 284  
 185 TGGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 244  
 285 CTGGTGTGACATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 344  
 245 CTGGTGTGACATTCATTCATTCATTCATTCATTCATTCATTCATTCATTCATTC 304  
 345 AGGATCACAAAATGAATTTTGTGATGCTTACGAAATCTCCACTCCAGTTTATC 404  
 305 AGGATCACAAAATGAATTTTGTGATGCTTACGAAATCTCCACTCCAGTTTATC 364  
 405 AGCTGGCGCAGCAGATGAAGAGGCGTGAAGCTGCTGTATGAATGAATGCTCTGCTG 464  
 365 AGCTGGCGCAGCAGATGAAGAGGCGTGAAGCTGCTGTATGAAGCTGCTCTGCTG 424  
 465 GGTCCCTACCAAGAAAGTGGCTGTAACTGCTTCAGTGAAGGAGGAGTCACTGCTAC 524  
 425 GGCCTTCACCAAGAAAGTGGCTGTAACTGCTTCAGTGAAGGAGGAGTCACTGCTAC 484  
 525 TACTGTGAGATTCAGATTCCTCCCACTGCGAAGAGGTTGATGCGCATGCT 584  
 485 TACTGTGAGATTCAGATTCCTCCCACTGCGAAGAGGTTGATGCGCATGCT 544  
 585 GTGAGCAGATTCAGATTCCTCCCACTGCGAAGAGGTTGATGCGCATGCT 644  
 545 GAGAGCGCGATTCAGATTCCTCCCACTGCGAAGAGGTTGATGCGCATGCT 604  
 645 TCTGTGTGAGCTTCCTCCCACTGCGAAGAGGTTGATGCGCATGCT 704  
 605 TCAATGTGTGAGCTTCCTCCCACTGCGAAGAGGTTGATGCGCATGCT 664  
 705 AGTTTGTGAGCTTCCTCCCACTGCGAAGAGGTTGATGCGCATGCT 764  
 665 AGTTTGTGAGCTTCCTCCCACTGCGAAGAGGTTGATGCGCATGCT 724  
 765 AACAGTCCCTACCGGCGCATGCGCTGCGAGTGGATCTGCGGGGAGGAGCGCGAATCT 824

Db 725 GACAGCCCTTACCCCGCTACCTGCGGAGGCGGCGGAGGAGGCGGAGTCA 784  
 QY 825 GTGCTGAGCTTCACTTCCGAACCTTGTGATGCTCTCTGTATGATGCAATGCACTGAC 884  
 Db 785 GTGCTGAGCTTCACTTCCGAACCTTGTGATGCTCTCTGTATGATGCAATGCACTGAC 844  
 QY 885 CTGGTACCGGTATGATGATGCTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 944  
 Db 845 CTGGTACCGGTATGATGATGCTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 904  
 QY 945 GGCACCTTCTGACCTTCTTCAAACTGACTTCTCTCTCTCCGAGAACTTCTCTGTC 1004  
 Db 905 GGCACCTTCTGACCTTCTTCAAACTGACTTCTCTCTCTCCGAGAACTTCTCTGTC 964  
 QY 1005 AGCCTATTAACCAATCTGACCGGCGACATCTGCTTGGAGCCACTTCTCTGACCTG 1064  
 Db 965 ACACATATTAACCAACCTGACCGGCGACATCTGCTTGGAGCCACTTCTCTGACCTG 1024  
 QY 1065 CCCAAGATGAGCAGCTGTGGCGGCTTTTGTAGTGAACCCCAAGGAGCATTTAGCAGCC 1124  
 Db 1025 CTTAGGATGAGCAGCTGTGGAGGCGCTTACGTAAAGCCCAAGGAGCATTTAGCAGCC 1084  
 QY 1125 TACTATCCAGGCGCATCTACCGCCCAACATCACTGCAATGAAATCAAGGTGCCAAC 1184  
 Db 1085 TACTATCCAGGCGCATCTACCGCCCAACATCACTGCAATGAAATCAAGGTGCCAAC 1144  
 QY 1185 AACCGGAAGTGAAGGTGGGCTTCAAACTTCTTCTCTGTGAGACCCCAAGTCAAGT 1244  
 Db 1145 AACCGGAAGTGAAGGTGGGCTTCAAACTTCTTCTCTGTGAGACCCCAAGTCAAGT 1204  
 QY 1245 GGCCTCTGACCAAGGAGTATGAGATCAACGGGAGAGAGTACTCGGTTGAGGAGTCC 1304  
 Db 1205 GGCACCTGCGCCCAAGGAGTATGAGATCAACGGGAGAGAGTACTCGGTTGAGGAGTCC 1264  
 QY 1305 CAGTTTGTGTGAGAGCAAGCAAGCAAGCAAGTAACTGCTCAATCTGATCACTG 1364  
 Db 1265 CAGTTTGTGTGAGAGCAAGCAAGCAAGCAAGTAACTGCTCAATCTGATCACTG 1324  
 QY 1365 TACAGGAGACCGGAGTCTTCTGAGTAACTCTCTCTGAGTCCAGACCGGTTGCCA 1424  
 Db 1325 TACAGGAGACCGGAGTCTTCTGAGTAACTCTCTCTGAGTCCAGACCGGTTGCCA 1384  
 QY 1425 GGGATGTTCAATGTGCAAGCTGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1484  
 Db 1385 GGGAGTTCACAGTTCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1444  
 QY 1485 GCAAGCTCCCGGATTAATGATGAGGCTTAATGCTCCAGATGCAATGCCACAGTTC 1544  
 Db 1445 GCGGATGCAAGGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1504  
 QY 1545 AGTGCAGAAACCAATTCGCAAGGCGCTCTTCTGAGTCTGAGACAGTGTCAAGACTGT 1604  
 Db 1505 AGTGCAGAAACCAATTCGCAAGGCGCTCTTCTGAGTCTGAGACAGTGTCAAGACTGT 1564  
 QY 1605 GGGAGCGAAGTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1664  
 Db 1565 GGAAGCAACAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1624  
 QY 1665 GGGAAAGTGTCTCTCAAGCAGCAAGATGTAATGGAAGCAACTGTGAGATGGTCT 1724  
 Db 1625 GGGAAAGTGTCTCTCAAGCAGCAAGATGTAATGGAAGCAACTGTGAGATGGTCT 1684  
 QY 1725 GACGAGGCTCATGAGAGGAGTGAATGCTCTCTGCAACCAATTAATCACTGCTGCT 1784  
 Db 1685 GACGAGGCTCATGAGAGGAGTGAATGCTCTCTGCAACCAATTAATCACTGCTGCT 1744  
 QY 1785 CAAATGAGCTCTGTCTGAGCAAGGAGCAACCTGATGTGAGGAGCAAGGAGTGTAGC 1844  
 Db 1745 CTCAATGAGCTCTGTCTGAGCAAGGAGCAACCTGATGTGAGGAGCAAGGAGTGTAGC 1804  
 QY 1845 GATGCTCCATGAGAAATCTGAGCTGTGGCTGCGATCTTTTACCAACAGGCTGCG 1904  
 Db 1805 GACGCTCATGATGAGAAAGTGTGAGCTGTGGCTGCGATCTTTTACCAACAGGCTGCT 1864

QY 1905 GTGTTGTGTCACGATGCGGACGAGGCGGAGTGCCCTTGGCAGGTGAGCTTCCAGCC 1964  
 DB 1865 GTTGTGGGGGACGATGCGATGAGGGCGAGTGCCCTTGGAGGTGAGCCCTGCAATGCT 1924  
 QY 1965 CTGGGCGAGGGGCACTTGTTGGGGGCTGCGCATCTCTCTGACTGCGGTCTGTGA 2024  
 DB 1925 CTGGGCGAGGGGCACTTGCGGTCTTCTCTCATCTCTCCCACTGGCTGTCTGCG 1984  
 QY 2025 GCTCATTTGCTTTCAGATGACAAATTTCAAGTACTCAGACTCAACGATTTGAGCGCC 2084  
 DB 1985 GCAACACTGCTACATCATGACAGAGATTCAAGTACTCAGACCCACGACATGAGCGCC 2044  
 QY 2085 TTCTCTGGGTCTGTGACACGAGCAAGCGAGTCTTGGGGTGGAGAGCTGAGAGCT 2144  
 DB 2045 TTCTCTGGGTCTGTGACACGAGCGAGCGAGCGCCCTGGGGTGGAGAGCGAGGCTC 2104  
 QY 2145 AAACGATATCATCAACCAACCTTCTCAATGATTTCACTTCACTATGACATCGCTTG 2204  
 DB 2105 AAGCGCATCTCTCCCACTCCCTTCTCATGACTTCACTTCACTTCACTATGACATCGGCTG 2164  
 QY 2205 CTGGAGCTGGAGAAAGTCGATGAGTACAGCACCGTCGTGCGCCCATCTGCTGCTGAT 2264  
 DB 2165 CTGGAGCTGGAGAAACCGGACAGTACAGCTCCATGATGTCGCGCCCATCTGCTGCGGAC 2224  
 QY 2265 GCTAACCATATCTTCTCTGTGCGCAAGGCGCATCTGGGTCAAGGCTGGGGGCAACAAAA 2324  
 DB 2225 GCTTCCCATCTTCTCTGTGCGGCAAGGCGCATCTGGGTCAAGGCTGGGGGCAACACCAG 2284  
 QY 2325 GAGGAGATACCGGAGCGTGTCTGTCAGAAAGGTGAGATCCGTTTCATCAACCAAGACC 2384  
 DB 2285 TAGGAGAGCATCTGGCGGCTGATCTCTGCAAAAGGTGAGATCCGCGTCATCAACCAAGACC 2344  
 QY 2385 ACCGTGAGACCTCATGCGCGACAGATACACCCCAAGATGATGTGTGGTTTCTC 2444  
 DB 2345 ACCGTGAGAACCTCTGCGCGACAGATACACCGCGCATGATGTGTGGGCTTCTC 2404  
 QY 2445 AGTGGGGGTGTGACCTCTGCGCAAGGCGCATCTGGGTGAGGCGGCTGCAAGGCGGAGAA 2504  
 DB 2405 AGCGGCGGCTGTGACCTCTGCGCAAGGCGATCTCGGGGAGCCCTGTCAGGCTGAGGCG 2464  
 QY 2505 GATGGGCGAATGTTCCAGGCTGTGTGTGAGTGTGGGTGAGAGCTCGCTCAAGAGAAC 2564  
 DB 2465 GATGGGCGGATCTTCCAGGCGGCTGTGTGAGTGTGGGAGAGCGCTCGCTCAAGAGAAC 2524  
 QY 2565 AAGCCAGCGCTGTACACAAAGGCTCTGTGATTTGGGACTGTGATCAAAAGAGCACTGGG 2624  
 DB 2525 AAGCCAGCGCTGTACACAAAGGCTCTGTGATTTGGGACTGTGATCAAAAGAGCACTGGG 2584  
 QY 2625 GTATAGCAGATGACAGACGACGACCAACCAACCAAGGATGCGGACATGACCA 2684  
 DB 2585 GTATAGGAGGCGGCGG--GCCACCAATGTGTACACTGCGGGGCGACCCATCTGTCAAC 2641  
 QY 2685 CCTGGATACAGAGAGAGACACTGACGACATTTATGTGTGGGCTTCCCGCCCAACCA 2744  
 DB 2642 CCACTGTGCAAG--CCTGAGGCTGTGAGACTGGAACCGCTGACATGACAGCGCCC--CGAGA 2699  
 QY 2745 ACCCAGACTGTGAATCTGATCTTTAGACTAGAGT 2780  
 DB 2700 ACATACACTGTGAATCTCAAGGCTCAAAAT 2735  
 RESULT 13  
 AAL60793  
 ID AAL60793 standard; DNA; 3147 BP.  
 XX AAL60793;  
 AC  
 XX  
 DT 03-SEP-2003 (first entry)  
 XX  
 DE Human membrane-type serine protease MTSpl protease domain DNA.  
 XX  
 KM Serine protease 17; CVP17; tumour; cancer; antisense therapy; prostate;

KM breast; cervix; lung; ovary; colon; gene therapy; human; enzyme; MTSPl;  
 KM membrane-type serine protease; matrixase; gene; ds.  
 OS Homo sapiens.  
 XX  
 XX  
 FT Key Location/Qualifiers  
 FT CDS 1865..2590  
 FT /tag= a  
 FT /product= "Human matrixase protease domain"  
 PN MO2003044179-A2.  
 PD 30-MAY-2003.  
 XX  
 PF 20-NOV-2002; 2002WO-US037626.  
 XX  
 PR 20-NOV-2001; 2001US-0332015P.  
 XX  
 PA (CORV-) CORVAS INT INC.  
 PI Madison EL, Ong EO;  
 XX  
 DR WPI; 2003-449816/42.  
 DR P-PSDB; AAO30147.  
 XX  
 PT New substantially purified serine protease 17 polypeptide and encoding  
 PT nucleic acid, useful for diagnosing and treating tumor conditions and/or  
 PT cancer, particularly of the breast, cervix, prostate, lung, ovary or  
 PT colon.  
 XX  
 PS Disclosure; Page 177-179; 189pp; English.  
 PS  
 XX  
 CC The invention relates to serine protease 17 polypeptide designated CVP17  
 CC and its corresponding nucleic acid sequence. The invention also relates  
 CC to a method using CVP17 protein to identify compounds that modulate its  
 CC protease activity. The method is useful for preventing, diagnosing and  
 CC treating disorders related to the serine protease 17 activity, such as  
 CC tumour conditions and/or cancer, particularly of the breast, prostate,  
 CC cervix, lung, ovary or colon. CVP17 DNA is used in gene therapy and in  
 CC antisense therapy. The present sequence is human membrane-type serine  
 CC protease MTSPl (also called matrixase) protease domain DNA. This  
 CC sequence is used to illustrate the method of the invention  
 CC  
 XX  
 SQ Sequence 3147 BP; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other;  
 Query Match 60.6%; Score 1883.2; DB 7; Length 3147;  
 Best Local Similarity 81.2%; Pred. No. 0;  
 Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;  
 QY 45 GATGGAACCGCAAAACCATGGGTAGCAATCGGGGCGGCAAGCGGAGGGGCTTCAG 104  
 DB 5 GAGCGGCTCGGGGTACCATGAGGAGCATCGGCGCGCAAGGCGGAGGGGCGCGAAG 64  
 QY 105 GACTTGGCGCGGAGCTCAAGTACAACTCCCGGTAGAGAAATGATGGCTTGGAGAG 164  
 DB 65 GACTTGGCGCGGAGCTCAAGTACAACTCCCGGACGAGAAAGTGAATGGCTTGGAGAG 124  
 QY 165 GGTGTGAGTCTCGCTGCGGAGCAATGCGCAAGAAAGTGGAGAGCGAGGCGCCAGGCGC 224  
 DB 125 GCGTGTGAGTCTCGCTGCGGAGCAATGCGGAGTGAATGGCTTGGAGAGAG 184  
 QY 225 TGGGTGTGTGTGTGGGCAAGTCTGTTCACTTCTCTTGTCTCTCCCTATGAGCTTG 284  
 DB 185 TGGGTGTGTGTGTGGGCAAGTCTGTTCACTTCTCTTGTCTCTCCCTATGAGCTTG 244  
 QY 285 CTGCTGTGCACTTCCATTAACGAAATGTGGGGTCAAAAAGTTCATAGGCGCATCTG 344  
 DB 245 CTGCTGTGCACTTTCATTAACGAAATGTGGGGTCAAAAAGTTCATAGGCGCATCTG 304  
 QY 345 AGATCACAAAATGAGATCTTCTGATGCGATGAGAACTCCACCTCCACAGATTATC 404  
 DB 305 AGATCACAAAATGAGATTTTGTGATGCTACGAGAACTCCACCTCCACAGATTATC 364



QY 405 AGCTGCGCAGCCAGTGAAGAGGCGGTGAAGCTGCTGTACATGAATGCTCTGTCTG 464  
DB 365 AGCTGCGCAGCAGAGTGAAGAGGCGGTGAAGCTGCTGTACATGAATGCTCTGTCTG 424  
QY 465 GGTGCTCTACCAAGAAAGTGGCTGTAACTGCTTCAAGTAGGGCAGTGTATGCGCTAC 524  
DB 425 GGGCCCTACCAAGAAAGTGGCTGTAACTGCTTCAAGTAGGGCAGTGTATGCGCTAC 484  
QY 525 TACTGTCAAGATTCAAGATCCCGCCACACTGTGCAAGAAAGTGTGATCGCGCATGGCT 584  
DB 485 TACTGTCAAGATTCAAGATCCCGCCACACTGTGCAAGAAAGTGTGATCGCGCATGGCT 544  
QY 585 GTGAGAGAGTGTGAATCATTTGCAACCCCGGAGAGCGGCACTGAAATCTTCTGTACCA 644  
DB 545 GAGAGAGAGTGTGAATCATTTGCAACCCCGGAGAGCGGCACTGAAATCTTCTGTACCA 604  
QY 645 TCTGTGTGGCTTCTTCCCAATTGACCCAGAAATGTGTGAGAGAGTGTGAGAGAGTGTG 704  
DB 605 TCAAGTGTGGCTTCTTCCCAATTGACCCAGAAATGTGTGAGAGAGTGTGAGAGAGTGTG 664  
QY 705 AGTTTGTGGCTTGTATGCTGCAAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTG 764  
DB 665 AGTTTGTGGCTTGTATGCTGCAAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTG 724  
QY 765 AACAGTCCCTACCCCGGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 824  
DB 725 GACAGTCCCTACCCCGGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 784  
QY 825 GTGCTGAGAGCTCACTTCCGAAAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 884  
DB 785 GTGCTGAGAGCTCACTTCCGAAAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 844  
QY 885 CTGCTGAGAGCTGTATATAGCTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 944  
DB 845 CTGCTGAGAGCTGTATATAGCTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 904  
QY 945 GGCAGCTTCTACCCCTGCAACACCTGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 1004  
DB 905 GGCAGCTTCTACCCCTGCAACACCTGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 964  
QY 1005 AGCTGATTAACCAATATGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1064  
DB 965 AACTGATTAACCAATATGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1024  
QY 1065 CCCAAGATGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1124  
DB 1025 CCTAGAGATGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1084  
QY 1125 TACTATCAAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 1184  
DB 1085 TACTATCAAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 1144  
QY 1185 AACCGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1244  
DB 1145 AACCGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1204  
QY 1245 GGCCTCTGCAACAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1304  
DB 1205 GGCCTCTGCAACAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1264  
QY 1305 CAGTTGTGAGAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1364  
DB 1265 CAGTTGTGAGAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1324  
QY 1365 TACACGAGAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1424  
DB 1325 TACACGAGAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1384  
QY 1425 GGGAGTGTGAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1484  
DB 1385 GGGAGTGTGAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1444  
QY 1485 GCAAGCTGCGGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 1544

DB 1445 GCCAGCTGACCGACCAACAGCATGAGCTCAACTGACAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1504  
QY 1545 ACGTGAAGAAACAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 1604  
DB 1505 ACGTGAAGAAACAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 1564  
QY 1605 GGGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1664  
DB 1565 GGGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1624  
QY 1665 GGGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1724  
DB 1625 GGGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1684  
QY 1725 GAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1784  
DB 1685 GAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1744  
QY 1785 CAAATGAGCTGTGCTGAGCAAGGAGCAACCTGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 1844  
DB 1745 CTCAATGAGCTGTGCTGAGCAAGGAGCAACCTGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 1804  
QY 1845 GATGCTGCTGATGAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1904  
DB 1805 GATGCTGCTGATGAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 1864  
QY 1905 GTGCTGTGAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 1964  
DB 1865 GTGCTGTGAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 1924  
QY 1965 CTGAGCAGAGGAGCACTGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 2024  
DB 1925 CTGAGCAGAGGAGCACTGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 1984  
QY 2025 GCTCATTTGCTTCAAGATGAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 2084  
DB 1985 GCTCATTTGCTTCAAGATGAGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 2044  
QY 2085 TTCTGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 2144  
DB 2045 TTCTGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 2104  
QY 2145 AAACGATCATCAACCACTTCTTCAATGATTTCACTTCAATGATTTCACTTCAATGATTTCACTTCAAT 2204  
DB 2105 AAACGATCATCAACCACTTCTTCAATGATTTCACTTCAATGATTTCACTTCAATGATTTCACTTCAAT 2164  
QY 2205 CTGAGCTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 2264  
DB 2165 CTGAGCTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 2224  
QY 2265 GCTATCCCATTTCTTCTTCTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 2324  
DB 2225 GCTATCCCATTTCTTCTTCTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 2284  
QY 2325 GAGGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 2384  
DB 2285 TATGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 2344  
QY 2385 AACTGTGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 2444  
DB 2345 AACTGTGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 2404  
QY 2445 AGTGGAGTGTGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 2504  
DB 2405 AGTGGAGTGTGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAG 2464  
QY 2505 GATGGAGTGTGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 2564  
DB 2465 GATGGAGTGTGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 2524  
QY 2565 AAGCGAGCTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAGAGTGTGAGAG 2624

Db 2525 AAGCCAGGCGGTGTACCAAGGCTCCTCTGTTTGGGACTGTGATCAAGAGAACTGGG 2584  
 QY 2625 GTATAGACACATGACAGACAGCCGACCAACACCCAGGGATGCCGACATGCACA 2684  
 Db 2585 GTATAGGCGCCGGG--GCCACCCCAATGTGTACCTGGCGGGGCCACCCATGTCTCAC 2641  
 QY 2685 CCTGTATACAGAGAGAGAACTGACGACATTATGCTGTGGCTCCCCCCCCCAACACA 2744  
 Db 2642 CCAAGTGTGACG-CCTGCAGGCTGTGAGACTGACCGCTGACTGACACGAGCCCC-CCAGA 2699  
 QY 2745 ACCCAGACTGTGAACTGACATCCTTAGAGACTCAGAGT 2780  
 Db 2700 ACATACACTGTGAACTCAATCTCCAGGGCTCCAAAT 2735

RESULT 14  
 AAL60792  
 ID AAL60792 standard; DNA; 3147 BP.  
 XX AC AAL60792;  
 XX DT 03-SEP-2003 (first entry)  
 XX DE Human membrane-type serine protease MTSPI DNA.  
 XX KM Serine protease 17, CVPSP17; tumour; cancer; antisenese therapy; prostate;  
 KW breast; cervix; lung; ovary; colon; gene therapy; human; enzyme; MTSPI;  
 XX membrane-type serine protease; matrilipase; gene; ds.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH CDS 23..2599  
 FT /\*tag= a  
 FT /product= "Human matrilipase"  
 XX  
 XX MO2003044179-A2.  
 XX  
 XX 30-MAY-2003.  
 XX PD  
 XX 20-NOV-2002; 2002WO-US037626.  
 XX PF  
 XX 20-NOV-2001; 2001US-0332015P.  
 XX PR  
 XX (CORV-) CORVAS INT INC.  
 XX PA  
 XX Madison EL, Ong EO;  
 XX PI  
 XX WPI; 2003-449816/42.  
 XX DR P-PSDB; AAO30146.  
 XX  
 XX New substantially purified serine protease 17 polypeptide and encoding  
 PT nucleic acid, useful for diagnosing and treating tumor conditions and/or  
 PT cancer, particularly of the breast, cervix, prostate, lung, ovary or  
 PT colon.  
 XX  
 XX Disclosure; Page 171-175; 189pp; English.  
 PS  
 XX The invention relates to serine protease 17 polypeptide designated CVPSP17  
 CC and its corresponding nucleic acid sequence. The invention also relates  
 CC to a method using CVPSP17 protein to identify compounds that modulate its  
 CC protease activity. The method is useful for preventing, diagnosing and  
 CC treating disorders related to the serine protease 17 activity, such as  
 CC tumour conditions and/or cancer, particularly of the breast, prostate,  
 CC cervix, lung, ovary or colon. CVPSP17 DNA is used in gene therapy and in  
 CC antisenese therapy. The present sequence is human membrane-type serine  
 CC protease MTSPI (also called matrilipase) DNA. This sequence is used to  
 CC illustrate the method of the invention  
 CC  
 XX Sequence 3147 BP; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other;  
 SQ  
 Query Match 60.6%; Score 1883.2; DB 7; Length 3147;  
 Best Local Similarity 81.2%; Pred. No. 0;

Matches 2223; Conservative 0; Mismatches 508; Indels 5; Gaps 3;  
 QY 45 GATTCGACCGCCAAAACCATGGGTATGCAATCGGCGCCGGAAGGCGCGGAGCTCTCAG 104  
 Db 5 GAGCGGCTTCGGGGATGACATGGGAGACGATCGGCCCCGACAGGGCCGAGGGGCCGAGG 64  
 QY 105 GACTTCGGCGCGGGACTCAAGTACAACTCCCGGCTTGAAGAACATGATGCTTTGAGAG 164  
 Db 65 GACTTCGGCGCGGGACTCAAGTACAACTCCCGGACGAGAAATGATGCTTTGAGAGAA 124  
 QY 165 GGTGTGAGATTCTGCTCCGGAACATGCGCAAGAAAGTGAAGGAGGAGGAGCCGAGCGC 224  
 Db 125 GCGGTGAAGTTCCTGCAATCAACAGTCMAAGAGTGAAGAGATGAGCCGCGAGGCGC 184  
 QY 225 TGGGTGATGTGTGTCAGAGTGTCTGAGTTCAGTTCCTCTTGTCTCTCCATGAGTGTG 284  
 Db 185 TGGGTGATGTGTGTCAGAGTGTCTGAGTTCAGTTCCTCTTGTCTCTCCATGAGTGTG 244  
 QY 285 CTGTGTGGACTTTCATTAATGGAATGTGGGATTCAAAAGTCTTCAATGGCCATCTG 344  
 Db 245 CTGTGTGGACTTTCATTAATGGAATGTGGGATTCAGAGTCTTCAATGGCCATCTG 304  
 QY 345 AGGATCACAAAATGAGATCTTTCTGATGGGTATGAGAACTCCACTCCACAGAGTTATC 404  
 Db 305 AGGATCACAAAATGAGATTTTGTGATGCTTACAGAACTCCACTCCACTCCAGTTTGA 364  
 QY 405 AGCCTGGCCAGCCAGGTGAAGAGAGCGCTGAAGCTGCTGTACATGAAGTCCCTGCTG 464  
 Db 365 AGCCTGGCCAGCCAGGTGAAGAGAGCGCTGAAGCTGCTGTACATGAAGTCCCTGCTG 424  
 QY 465 GGTCTTACCAAGAAAGTGGCTGTAACTGCTTCACTGATGAGGGAGTGCATCGCTAC 524  
 Db 425 GGCCTTACCAAGAAAGTGGCTGTAACTGCTTCACTGATGAGGGAGTGCATCGCTAC 484  
 QY 525 TACTGTCAAGTTCAGCAATCCGCCACACCTGGCAAGAGTGTATGGCGCATGCT 584  
 Db 485 TACTGTCAAGTTCAGCAATCCGCCACACCTGGCAAGAGTGTATGGCGCATGCT 544  
 QY 585 GTGAGCGAGTGTGAATTCGACACCCCGAGCAAGGAGCTGAATCTTGTGCTTACA 644  
 Db 545 GAGAGCGCGTGAATTCGACACCCCGAGCGAGTGTGAATCTTGTGCTTACA 604  
 QY 645 TCTGTGTGCTTCCCATTTGACCCCAAGATGCTGAGAGGATCTCAGCAAGCTGC 704  
 Db 605 TCAGTGTGCTTCCCATTTGACCCCAAGATGCTGAGAGGATCTCAGCAAGCTGC 664  
 QY 705 AGTTTGCCTGATGCGCAATGAGAGAGTGAACAGCTTCACTACCCCTGCTGCC 764  
 Db 665 AGTTTGCCTGATGCGCAATGAGAGAGTGAACAGCTTCACTACCCCTGCTGCC 724  
 QY 765 AACAGTCCCTAACCCGCGGATGCCGCTGCAAGTGGTCTGCGGGGGAGCGCGACTCT 824  
 Db 725 GACAGCCCTTACCCCGCTCAATGCCGCTGCAAGTGGTCTGCGGGGGAGCGCGACTCA 784  
 QY 825 GTGCTGAGCTTCACTTCCGAAGCTTTGATGTGCTCTCTGTATGAGCAATGGCGATAC 884  
 Db 785 GTGCTGAGCTTCACTTCCGAAGCTTTGATGTGCTCTCTGTATGAGCAATGGCGATAC 844  
 QY 885 CTGTGACCCGTGTATGATAGCTGAGCCCATGGAACCCCAAGCTGTGTGTGCTGTGT 944  
 Db 845 CTGTGACCCGTGTATGATAGCTGAGCCCATGAGCCCATGAGCCCATGAGTGTGTGT 904  
 QY 945 GGCACCTTCTACCCCTCTTACACCTGACTTCTCTCTCTCCAGAAAGTCTTCTTGTG 1004  
 Db 905 GGCACCTTCTACCCCTCTTACACCTGACTTCTCTCTCTCCAGAAAGTCTTCTTGTG 964  
 QY 1005 AGCTGATTAACCAATCTGACCGGAGAAATCTGCTTTGAGGCGCACTTCTCCAGCTG 1064  
 Db 965 AACTGATTAACCAACACTGAGGCGGAGTCCCGCTTGAAGGCGCACTTCTCCAGCTG 1024  
 QY 1065 CCAAGATGAGAGAGTGTGGCGGCTTTTGAAGTGAACCAAGGAGCAATTAGACGCCCC 1124  
 Db 1025 CCTAGATGAGAGAGTGTGGCGGCTTTTGAAGTGAACCAAGGAGCAATTAGACGCCCC 1084

QY 1125 TACTATCCAGGCGCTACCTCCGCAAGATCACTGCACTATGGAAATATCAAGTCCCAAC 1184  
 DB 1085 TACTATCCAGGCGCTACCTCCGCAAGATCACTGCACTATGGAAATATCAAGTCCCAAC 1144  
 QY 1185 AACCGAAGCTGGAAGTGGCTTCAAACTCTTCTATCTGAGGACCCCAAGTACAGT 1244  
 DB 1145 AACCGAAGCTGGAAGTGGCTTCAAACTCTTCTATCTGAGGACCCCAAGTACAGT 1204  
 QY 1245 GGTCTCTGCAACCAAGACTATGTGGAGATCAACGGGGAGAAGTACTGCGGTGAGAGTCC 1304  
 DB 1205 GGGACCTGCCCCAAGGACTACGTGGAGATCAATGGGAGAAATATCTGGAGAGAGTCC 1264  
 QY 1305 CAGTTTGTGGTGAAGGCAACAGAGAAATTAAGTCACTTCCATTTGTATCACTCG 1364  
 DB 1265 CAGTTTGTGGTGAAGGCAACAGAGAAATTAAGTCACTTCCATTTGTATCACTCG 1324  
 QY 1365 TACACGCAACCGGGTCTTCTAGCTGAGTACTCTCCAGCACTCCCAACCGGTCGCA 1424  
 DB 1325 TACACGCAACCGGGTCTTCTAGCTGAGTACTCTCCAGCACTCCCAACCGGTCG 1384  
 QY 1425 GGGATGTTCACTGTCGAAGCTGACCGTGCATCCGAAGGAATGCGCTGCGACGCTGG 1484  
 DB 1385 GGGCACTTCACTGTCGCGCAACGGGCGGTGTATCCGGAAGGAGCTGGCGTGTATGGTGG 1444  
 QY 1485 GCAGACTGCGCGGATTAATGATGAGCGCTTACTGCGCAATGCACTCCCAACCGCTTC 1544  
 DB 1445 GCGCACTGCAACCAACAGCGATGAGCTCACTGAGTGGCACTCCCGGCAACCGATTC 1504  
 QY 1545 ACGTGCMAAAACCAAGTCTGCAAGCCCTCTTCTGGGCTGTGAGCAAGTGCMAAGCTGG 1604  
 DB 1505 ACGTGCMAAAACCAAGTCTGCAAGCCCTCTTCTGGGCTGTGAGCAAGTGCMAAGCTGG 1564  
 QY 1605 GGGGACGGAAGTGAACGAGAGGGCTGACGCTGTCTGCGGAGTTTCAAGTGTCCAAAT 1664  
 DB 1565 GGGGACGGAAGTGAACGAGAGGGCTGACGCTGTCTGCGGAGTTTCAAGTGTCCAAAT 1624  
 QY 1665 GGGAAAGTCTGCTCTGAGGCGCAAGAGTGAATGGGAAGCAACTGTGAGAGTGGCTCT 1724  
 DB 1625 GGGAAAGTCTGCTCTGAGGCGCAAGAGTGAATGGGAAGCAACTGTGAGAGTGGCTCT 1684  
 QY 1725 GACGAGCTTCACTGATGACGCTGATGCTGCTCTTCTGACCAAAATATCACTACCGCTGC 1784  
 DB 1685 GACGAGCTTCACTGATGACGCTGATGCTGCTCTTCTGACCAAAATATCACTACCGCTGC 1744  
 QY 1785 CAAATGCGCTCTGTCTGAGCAAGGGCAACCTGAGTGTGATGGGAAGACGACTGTAGC 1844  
 DB 1745 CTAATGCGCTCTGTCTGAGCAAGGGCAACCTGAGTGTGATGGGAAGACGACTGTAGC 1804  
 QY 1845 GATGCTCCGATGAGAAACTGTGATGCTGAGCTGCTCTTTCCTTTCCTTTCCTTTCCTTTC 1904  
 DB 1805 GATGCTCCGATGAGAAACTGTGATGCTGAGCTGCTCTTTCCTTTCCTTTCCTTTCCTTTC 1864  
 QY 1905 GTGTTGTGTCGACCAATGCGAGACGAGGGGAGTGGGCTCTGACAGTGGCTTCCAGCGCC 1964  
 DB 1865 GTGTTGTGTCGACCAATGCGAGTGGGAGTGGGCTCTGACAGTGGCTTCCAGCGCT 1924  
 QY 1965 CTGAGCTCAGGCTCACTTGTGTGGGCTCTGCTCACTCTCTCTGACTGAGTGTCTCTGA 2024  
 DB 1925 CTGAGCTCAGGCTCACTTGTGTGGGCTCTGCTCACTCTCTCTGACTGAGTGTCTCTGA 1984  
 QY 2025 GCTCATTTGCTTCAAGATGACAAATTTCAAGTACTCAAGTACTCAAGTGTGAGCGCC 2084  
 DB 1985 GCACTCTCTCATGATGAGACGAGATTCAGTACTCAAGTACTCAAGTGTGAGCGCC 2044  
 QY 2085 TTCTGGGTCTGTCGAGACCAAGACGAGTGGCTCTGAGGAGTGCAGAGCTGAAGCTC 2144  
 DB 2045 TTCTGGGTCTGTCGAGACCAAGACGAGTGGCTCTGAGGAGTGCAGAGCTGAAGCTC 2104  
 QY 2145 AAAGATATCATACCACTCTTCTTCAATGATTTCACTTCACTATGACATGCTGCTTG 2204  
 DB 2105 AAAGATATCATACCACTCTTCTTCAATGATTTCACTTCACTATGACATGCTGCTTG 2164

QY 2205 CTGAGACTGAGAAAGTGGTGAAGTACAGACCGCTGCGGCCCATCTGCTCTGAT 2264  
 DB 2165 CTGAGACTGAGAAAGTGGTGAAGTACAGACCGCTGCGGCCCATCTGCTCTGAT 2224  
 QY 2265 GCTACCATATCTTCTCTGTCGAGAGGCGCATCTGGGTCAACAGCTGGGGGACACAAAA 2324  
 DB 2225 GCTACCATATCTTCTCTGTCGAGAGGCGCATCTGGGTCAACAGCTGGGGGACACAAAA 2284  
 QY 2325 GAGGAGGTACCGGAGCGCTGATCTCTGACAGAGGAGTACCTGCTCATCAACCAACC 2384  
 DB 2285 TATGAGAGCACTGCGCGCTGATCTGCAAAAGGAGTACCTGCTCATCAACCAACC 2344  
 QY 2385 ACCTGAGAGCACTGATGCGGACAGATCAACCAAGTATGATGATGATGATGATGATGAT 2444  
 DB 2345 ACCTGAGAGCACTGATGCGGACAGATCAACCGCGCATGATGATGATGATGATGATGAT 2404  
 QY 2445 AGTGGGGGTGAGATCTCTGCAAGGATGATCTGATGATGATGATGATGATGATGATGAT 2504  
 DB 2405 AGTGGGGGTGAGATCTCTGCAAGGATGATCTGATGATGATGATGATGATGATGATGAT 2464  
 QY 2505 GATGGGCAATGTTTCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2564  
 DB 2465 GATGGGCAATGTTTCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 2524  
 QY 2565 AAGCCAGGCGGTGACAAAGGCTCCTGATGATGATGATGATGATGATGATGATGATGATGAT 2624  
 DB 2525 AAGCCAGGCGGTGACAAAGGCTCCTGATGATGATGATGATGATGATGATGATGATGATGAT 2584  
 QY 2625 GTATGACACATGAGACAGACAGCCGACCAAAACACCAACAGGAGTCCGACATGACAC 2684  
 DB 2585 GTATGACACATGAGACAGACAGCCGACCAAAACACCAACAGGAGTCCGACATGACAC 2644  
 QY 2685 CCGATACAGAGAGAGAACATGACATTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2744  
 DB 2642 CCGATGACAGC-CCTGACGCTGAGACTGACCGCTGACCTGACCAAGCGCC-CCAG 2699  
 QY 2745 ACCCAAGCTGTGAATGATCTCTTGAAGTCAAGT 2780  
 DB 2700 ACATACACTGTGAATGATCTCTTGAAGTCAAGT 2735

RESULT 15  
 AAD47225 strand; DNA; 3147 BP.  
 AAD47225;  
 AAD47225; (first entry)  
 24-FEB-2003  
 Human membrane-type serine protease 1 (MTSP1) DNA.  
 Human type II membrane-type serine protease 9; tumour; transgenic;  
 Human type II transmembrane serine protease; enzyme; gene therapy; MTSP9;  
 neoplastic disease; transgenic animal; membrane-type serine protease 1;  
 TTS1; MTSP1; matrilysin; gene; ds.  
 Homo sapiens.  
 Key Location/Qualifiers  
 CDS 23..2590  
 FT /tag=a  
 FT /product="Human MTSP1 protein"  
 MO20027267-A2.  
 03-OCT-2002.  
 27-MAR-2002; 2002WC-US009611.  
 27-MAR-2001; 2001US-0279228P.  
 15-MAY-2001; 2001US-0291501P.  
 (CORV-) CORVAS INT INC.

XX Madison EL, Ong EO;  
 XX WPI, 2003-018940/01.  
 DR P-PSDB; AAE29820.  
 XX  
 PT New substantially purified single or two-chain type II membrane-type  
 PT serine protease 9 (MTSP9) polypeptide, useful for monitoring tumor  
 PT progression, inhibiting tumor initiation, or treating a malignant or pre-  
 PT malignant condition.  
 XX  
 PS Disclosure; Page 179-183; 139pp; English.  
 XX  
 CC The invention relates to type II membrane-type serine protease 9 (MTSP9)  
 CC polypeptides and polynucleotides. MTSP belongs to type II transmembrane  
 CC serine protease (TSP) family. Sequences of the invention and their  
 CC antibodies are useful for diagnosing, treating or preventing neoplastic  
 CC disease in mammals. They are useful for monitoring tumor progression,  
 CC inhibiting tumor initiation, growth or progression or treating malignant  
 CC or pre-malignant conditions. Transgenic animals of the invention are  
 CC useful in animal models of tumor initiation, growth and/or progression  
 CC models. The invention is also useful in gene therapy. The present  
 CC sequence is human membrane-type serine protease I (MTSP1) DNA. MTSP1 also  
 CC referred as matrilipase is a member of the TSP family  
 XX  
 SQ Sequence 3147 BP; 654 A; 952 C; 958 G; 583 T; 0 U; 0 Other:  
 Query Match 60.6%; Score 1893.2; DB 7; Length 3147;  
 Best Local Similarity 81.2%; Pred. No. 0;  
 Matches 222; Conservative 0; Mismatches 508; Indels 5; Gaps 3;  
 QY 45 GATTCGACCCGCCAAACCATGGGTAGCAATCGGGGGCGCAAGGCGGAGGGGGCTCTCAG 104  
 DB 5 GAGGGGCTCGGGGTACCATGAGGAGCATGGGCCCGCAAGGGCGGAGGGGCCCAAG 64  
 QY 105 GACTTCGGCGCGGGACTCAAGTACAACTCCCGCTAGAGAAATGAAATGGCTTTGAGAG 164  
 DB 65 GACTTCGGCGCGGGACTCAAGTACAACTCCCGCGCAAGAAATGAAATGGCTTTGAGAG 124  
 QY 165 GGTGTGAGATTCTCTGCTCGGAGCAATGTCGAAAGAAATGAGAAAGCGAGGCCCGCGC 224  
 DB 125 GGCCTGAGATTCTCTGCGAGTCACAAAGTCMAAGAGGTGGAAAGCATGGGCCCGGGCGC 184  
 QY 225 TGGGTGTGCTGT 284  
 DB 185 TGGGT 244  
 QY 285 CTGGT 344  
 DB 245 CTGGT 304  
 QY 345 AGGATCAAAATGAAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 404  
 DB 305 AGGATCAAAATGAAATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 364  
 QY 405 AGGCTGT 464  
 DB 365 AGGCTGT 424  
 QY 465 GGTCTCTACACAAAGAGTCGGCTGTAACTGTCTTCACTGAGGGGCACTGTCTAC 524  
 DB 425 GGGCCCTTACCAAGAGAGTCGGCTGTAACTGTCTTCACTGAGGGGCACTGTCTAC 484  
 QY 525 TACTGT 584  
 DB 485 TACTGT 544  
 QY 585 GTGGAGGAGTTGTAAATTTGTCACCCCGAGAGCAAGGAGCTGAAATCTTGTGTGTAA 644  
 DB 545 GAGGAGGAGTTGTAAATTTGTCACCCCGAGAGCAAGGAGCTGAAATCTTGTGTGTAA 604  
 QY 645 TCTGT 704

DB 605 TCAGT 664  
 QY 705 AGTTTGT 764  
 DB 665 AGTTTGT 724  
 QY 765 AACAGT 824  
 DB 725 GACAGGCTTACCCGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 784  
 QY 825 GT 884  
 DB 785 GT 844  
 QY 885 CTGGT 944  
 DB 845 CTGGT 904  
 QY 945 GGCACCTTCTACCTCTCTCTCAACCTGACTTTCTCTCTCTCTCTCTCTCTCTCTCT 1004  
 DB 905 GGCACCTTCTCTCTCTCTCTCAACCTGACTTTCTCTCTCTCTCTCTCTCTCTCTCT 964  
 QY 1005 AGCTGT 1064  
 DB 965 ACAGT 1024  
 QY 1065 CCCAAGATGAGCAGCTGT 1124  
 DB 1025 CTGATGATGAGCAGCTGT 1084  
 QY 1125 TACTATTCAGGCTGT 1184  
 DB 1085 TACTATTCAGGCTGT 1144  
 QY 1185 AACCGAAGT 1244  
 DB 1145 AACCGAAGT 1204  
 QY 1245 GGTCTCTGT 1304  
 DB 1205 GGTCTCTGT 1264  
 QY 1305 CAGTTTGT 1364  
 DB 1265 CAGTTTGT 1324  
 QY 1365 TACACGAGACCGGGT 1424  
 DB 1325 TACACGAGACCGGGT 1384  
 QY 1425 GGT 1484  
 DB 1385 GGT 1444  
 QY 1485 GCAAGT 1544  
 DB 1445 GCAAGT 1504  
 QY 1545 ACCTGT 1604  
 DB 1505 ACCTGT 1564  
 QY 1605 GGGAGCGAAGT 1664  
 DB 1565 GGGAGCGAAGT 1624  
 QY 1665 GGGAGT 1724  
 DB 1625 GGGAGT 1684  
 QY 1725 GACGAGGCTTATGT 1784  
 DB 1685 GACGAGGCTTATGT 1744



BLANK  
PAGE

---